ROLE OF INSTITUTIONAL CREDIT IN INDIAN AGRICULTURAL PRODUCTION: A DETAILED TIME SERIES ANALYSIS

Amrita Bhattacharya

Southern Illinois University Carbondale

ABSTRACT

In this paper, I look at the relation between institutional credit and agricultural production through time series analysis. This analysis gave some expected yet not so desired results. Stationarity and Co-integration, two major characteristics of time series analysis and long run relationship between variables have been checked for the chosen variables, followed by identification of the stochastic process involved in each series. The series we’ve chosen for analyzing the ‘Effect Institutional Credit on Indian Agriculture’ are the following: Production of food grain and major commercial crop in India in between 1970-2008 and Institutional credit to agricultural sector over the same time period. The empirical results suggest that Indian agriculture can improve a lot if sufficient amount of credit is issued to agricultural sector and if the issued fund is used efficiently. Farmers have to depend on non-institutional credit sources. Besides, complex credit policies have also refrained the farmers from taking a step towards institutional sources. The results also show that dependence on monsoon results in fluctuation in agricultural production, which means farmers’ default rate increases in times of bad monsoon, so banks do not issue credit to them. In turn, lack of fund compels the farmers to stick to old production techniques, behavior of monsoon and non-institutional sources of credit charging high interest rates. So, productivity doesn’t rise significantly.

JEL Classification: O1, O13, Q1, Q18

Keywords: Institutional credit, Agricultural production, Stationarity, Co-integration

1. INTRODUCTION

Agriculture plays a crucial role in the development of the Indian economy. A large proportion of the population in India is rural based and depends on agriculture for a living. Enhanced and stable growth of the agriculture sector is important as it plays a vital role not only in generating purchasing power among the rural population by creating on-farm and off-farm employment opportunities but also through its contribution to price stability. Credit is a crucial factor in
agricultural production and in many cases may be a limiting factor in small holder agriculture. According to Miller (1977), credit provides the means for the temporary transfer of assets from an individual or organization to one which has not. Credit may be described as a facility extended from the lender to the borrower and is repayable at maturity, which may range from a few days to several years. For a credit transaction to be completed, the borrower must provide some evidence of debt obligation in return for the loan where the loan is based solely on good reputation, financial position of the borrower and trust. Credit can also be extended to the borrower in the form of assets possessed by the lender i.e. in cash (Miller 1977; Abayomi and Salami, 2008). A strong and efficient agricultural sector has the potential to enable a country feed its growing population, generate employment, earn foreign exchange and provide raw materials for industries. The vibrancy of the sector has a multiplier effect on any nation's socio-economic and industrial fabric, because of multifunctional nature. A number of studies such as Ansari, Gerasim and Mahdavinia (2009), and Salami, et al (2010) have documented the problems of the agricultural sector in Africa countries. Aside the problem of poor access to modern technology by the peasant farmers in the African countries, the major bane of agricultural development commonly identified by the above studies among others is low investment or finance. Credit plays a major role in the transformation of traditional agriculture into a modern large scale commercial type which enhances agricultural development. It is necessary for purchasing inputs needed for effective adoption of modern agricultural techniques. Many economists have identified the lack of basic assets major constraint to agricultural development (Abayomi and Salami, 2008). Oluwasanmi and Alao (1965) clearly stated the need for credit or the purchase of farm inputs such as improved seed varieties, breeds of livestock, fertilizers, insecticides, pesticides, modern implement, among others. They also stressed the suitability of terms of credit as a necessary condition for fostering agricultural development. Oyatoye (1981) averred that credit is a major factor necessary for technological transfer in traditional agriculture. According to her, given the availability of inputs needed to improve technology, how rapidly farmers would adopt improved technology depend on additional factors. She further identified efficient source of production credit as one of these additional factors. Oni (1987) opined that the peasant farmers do not possess enough resources to purchase these farm investments. He further stressed that it is necessary to supplement the farmer’s personal earnings to facilitate agricultural transformation. Hence the need for credit is universal. While it is needed by the less developed countries to increase productivity per farm worker and per hectare, the developed nations also need it to foster development (Jekayinfa, 1981; Abalu et al, 1981). Cole (2008) integrated theories of political budget cycles with theories of tactical electoral redistribution to test for political capture in a novel way. Studying banks in India, he found that government-owned bank lending tracks the electoral cycle, with agricultural credit increasing by 5-10 percentage points in an election year. There is significant cross-sectional targeting, with large increases in districts in
which the election is particularly close. This targeting does not occur in non-election years, or in private bank lending. He showed that capture is costly: elections affect loan repayment, and election year credit booms do not measurably affect agricultural output. Sreeram (2007) concluded that increased supply and administered pricing of credit help in the increase in agricultural productivity and the well-being of agriculturists as credit is a sub-component of the total investments made in agriculture. He also stated that the diversity in cropping patterns, holding sizes, productivity, regional variations make it difficult to establish a causality for agriculture or rural sector. In the last five decades, the Government’s objectives in agricultural policy and the instruments used to realize the objectives have changed from time to time, depending on both internal and external factors. Agricultural policies at the sectoral level can be further divided into supply side and demand side policies. The former includes those relating to land reform and land use, development and diffusion of new technologies, public investment in irrigation and rural infrastructure and agricultural price supports. The demand side policies on the other hand, include state interventions in agricultural markets as well as operation of public distribution systems. Such policies also have macro effects in terms of their impact on government budgets. Macro level policies include policies to strengthen agricultural and non-agricultural sector linkages and industrial policies that affect input supplies to agriculture and the supply of agricultural materials.

An important aspect that has emerged in last three decades is that the credit is not only obtained by the small and marginal farmers for survival but also by the large farmers for enhancing their income. But in India the overall thrust of the current policy regime assumes that credit is a critical input that affects agricultural or rural productivity and is important enough to establish causality with productivity. Therefore, in this backdrop I have undertaken the case study of determining any co-integration between the credit to agrarian sector and its production of major crops. An analysis of several sets of data for the same sequence of time periods is called multiple or multivariate time series analysis. The series chosen for analyzing the ‘Effect Institutional Credit on Indian Agriculture’ are the following: Production of food grain and major commercial crop in India in between 1970-2008 and Institutional credit to agricultural sector over the same time period. I wish to study the dynamics or temporal structure of the data by time series analysis. The paper is divided in 4 sections apart from introduction at the beginning. Section 2 gives a brief theoretical background and section 3 describes the fundamentals of time series analysis. Section 4 gives a brief description of the data. In section 5, I start the econometric analysis which is subdivided according to the progress of the analysis. Starting with the unit root test, followed by Correlogram analysis and stationary, identification of stochastic process and co integration analysis, section 5 ends with a summary of our findings from the econometric analysis. Section 6 concludes the analysis by explaining the economic backdrop of the findings.
2. THEORETICAL BACKGROUND

The importance of farm credit as a critical input to agriculture is reinforced by the unique role of Indian agriculture in the macroeconomic framework and its role in poverty alleviation. Agricultural policies in India have been reviewed from time to time to maintain pace with the changing requirements of the agriculture sector, which forms an important segment of the priority sector lending of scheduled commercial banks (SCBs). In India the need for affordable, sufficient and timely supply of institutional credit to agriculture has assumed critical importance. The demand for agricultural credit arises due to i) lack of simultaneity between the realization of income and act of expenditure; ii) lumpiness of investment in fixed capital formation; and iii) stochastic surges in capital needs and saving that accompany technological innovations. Recognizing the importance of agriculture sector in India’s development, the Government and the Reserve Bank of India (RBI) have played a vital role in creating a broad-based institutional framework for catering to the increasing credit requirements of the sector.

The trend in institutional agricultural credit from 1970 to 2008 is depicted in the diagram:

![Graph showing institutional agricultural credit trend from 1973-74 to 2005-06](image)

Three main factors that contribute to agricultural growth are increased use of agricultural inputs, technological change and technical efficiency. With savings being negligible among the small farmers, agricultural credit appears to be an essential input along with modern technology for higher productivity.
Hence, since independence, credit has been occupying an important place in the strategy for development of agriculture. The agricultural credit system of India consists of informal and formal sources of credit supply. The informal sources include friends, relatives, commission agents, traders, private moneylenders, etc. Three major channels for disbursement of formal credit include commercial banks, cooperatives and micro-finance institutions (MFI) covering the whole length and breadth of the country. A large number of formal institutional agencies like Cooperatives, Regional Rural Banks (RRBs), Scheduled Commercial Banks (SCBs), Non-Banking Financial Institutions (NBFIs), and Self-help Groups (SHGs), etc. are involved in meeting the short- and long-term needs of the farmers. Several initiatives have been taken to strengthen the institutional mechanism of rural credit system. The main objective of these initiatives was to improve farmers’ access to institutional credit. The major milestones in improving the rural credit are acceptance of Rural Credit Survey Committee Report (1954), nationalization of major commercial banks (1969 & 1980), establishment of RRBs (1975), establishment of National Bank for Agriculture and Rural Development (NABARD) (1982) and the financial sector reforms (1991 onwards), Special Agricultural Credit Plan (1994-95), launching of Kisan Credit Cards (KCCs) (1998-99), Doubling Agricultural Credit Plan within three years (2004), and Agricultural Debt Waiver and Debt Relief Scheme (2008). These initiatives had a positive impact on the flow of agricultural credit.

During the pre-green revolution period, from independence to 1964-1965, the agricultural sector grew at annual average of 2.7 per cent. This period saw a major policy thrust towards land reform and the development of irrigation. With the green revolution period from the mid-1960s to 1991, the agricultural sector grew at 3.2 per cent during 1965-1966 to 1975-1976, and at 3.1 per cent during 1976-1977 to 1991-1992. Acharya (1998) explains that the policy package for this period was substantial and consisted of: a) introduction of high-yielding varieties of wheat and rice by strengthening agricultural research and extension services, b) measures to increase the supply of agricultural inputs such as chemical fertilizers and pesticides, c) expansion of major and minor irrigation facilities, d) announcement of minimum support prices for major crops, government procurement of cereals for building buffer stocks and to meet public distribution needs, and e) the provision of agricultural credit on a priority basis. This period also witnessed a number of market intervention measures by the central and state Governments. The promotional measures relate to the development and regulation of primary markets in the nature of physical and institutional infrastructure at the first contact point for farmers to sell their surplus products.

Growth in the agriculture sector may well be judged by the increase in agricultural production over time. In economic terms, relative changes in prices of different crops also may effect substitution. In the Indian context, rice, wheat, cereals and pulses are the major food crops.
Oilseeds, sugarcane, cotton, jute & tobacco are the major cash crops. From 1970 to 2009 the overall trend in these food crops and cash crops are captured in the following diagram.
3. TIME SERIES ANALYSIS

A time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals. A time series is a collection of observations of well-defined data items obtained through repeated measurements over time. In other words, it is a sequence of numerical data in which each item is associated with a particular instant in time.

An analysis of single sequence of data is called univariate time series analysis.

An analysis of several set of data for the same sequence of time periods is called multivariate time series analysis.

Time series methods can be roughly divided into 2 types of methods: frequency-domain methods and time domain methods. However, in this project our approach will only be via time-domain methods. Time series analysis techniques may further be divided into parametric and non-parametric methods. The parametric approaches assume that the underlying stationary stochastic process has a certain structure which can be described using a small number of parameters (for example, using an autoregressive or moving average model). In these approaches, the task is to estimate the parameters of the model that describes the stochastic process. By contrast, non-parametric approaches explicitly estimate the covariance or the spectrum of the process without assuming that the process has any particular structure.

Broadly speaking, there are five approaches to economic forecasting based on the time series data:

1. Exponential Smoothing Methods
2. Single Equation Regression Model
3. Simultaneous Equation Regression Models
4. Autoregressive Integrated Moving Average Models
5. Vector Auto regression

A time series is a sequence of numerical data in which each item is associated with a particular instant of time. The basic assumption of time series analysis is that it has been generated by a stochastic process, i.e., each element of the series is drawn randomly from a probability distribution. Hence, it is a collection of random variable (Xₜ). Such a collection ordered in time is called Stochastic Process. If it is a continuous variable, it is denoted the random variable by X(t) and if t is a discrete variable, it is denoted them by Xₜ. The random variables are not independent in general. Furthermore, we have just a sample size 1 on each of the random variables. There is no way of getting another observation, so we are called a ‘single realization’. The two features
are dependence and lack of replication, compel us to specify some highly restrictive models for the statistical structure of the stochastic process.

1) Stationary:

One important class of stochastic process is that of stationary process. It guarantees that there are no fundamental changes in the structure of the process that would render prediction difficult or impossible. Corresponding to these we have the concept of stationary time series.

- **STRICT STATIONARITY:**

A time series is said to be strictly stationary if the joint distribution of any set of observations \( X(t_1), X(t_2), \ldots, X(t_n) \) is the joint distribution of \( X(t_1+k), X(t_2+k), \ldots, X(t_n+k) \) for all \( n \) and \( k \). Strict stationary holds for all values for \( n \) and a constant for all \( t \). Properties of strict stationary are:

- The mean \( \mu(t) = E(X_t) \).
- The variance \( \sigma^2(t) = \text{var}(X_t) \).
- The auto covariance function \( y(t_1, t_2) = \text{cov}(X_{t_1}, X_{t_2}) \).
- When \( t_1 = t_2 = t \), the autocovariance is just \( \sigma^2(t) \).

For a strictly stationary time series the distribution of \( X(t) \) is independent of \( t \). Thus it is not just the mean and variance which is constant but also all the higher order moments are independent of \( t \). So are all the higher order moments of joint distribution of any combinations of the variables \( X(t_1), X(t_2), \ldots \).

- **WEAK STATIONARITY:**

Thus, a time series is said to be weakly stationary if its mean is constant and its auto covariance function just depends on the difference \( (t_2 - t_1) \), which is called the lag. Hence, we can write the auto covariance function \( y(t_1, t_2) \) as \( y(k) \) where \( k = t_2 - t_1 \) the lag. So, the properties of the weak stationary are:

- The mean \( \mu(t) = E(x_t) \).
- The variance \( \sigma^2(t) = \text{var}(x_t) \).
- The auto covariance function \( y(k) = \text{cov}(X_{t_1}, X_{t_2}) \).

Since \( \text{var} X(t) = \text{var} X(t+k) = \sigma^2 = y(0) \), then we have the auto correlation coefficient \( \rho(k) \) at a lag \( k \) as \( \rho_k = y(k)/y(0) \). \( P_k \) is called the auto correlation function and will be abbreviated as auto correlation function. A plot of \( \rho(k) \) against \( k \) is called a correlogram.
2) Non-Stationary:

In time series analysis we do not confine ourselves to the analysis of stationary time series. In fact, most of the time series we encounter are nonstationary. A simple non-stationary time series model is \( X_t = \mu_t + e_t \), where mean \( \mu \) is a function of time and \( e \) is weakly stationary series. A time series is said to be nonstationary if its mean is a function of a time. So, mean is a linear or quadratic function of a \( t \).

Suppose, a stochastic process model is: \( x_t = \rho x_{t-1} + e_t \)

Where, \( \rho \) is a number between (-1) and (+1) and \( e_t \) is a sequence which is independent or uncorrelated identically distributed random variable with zero mean that is,

\[
\begin{align*}
\text{a)} & \quad \mathbb{E}(e_t) = 0 \\
\text{b)} & \quad \text{var}(e_t) = \sigma^2 e < \infty \text{ for all } t \\
\text{c)} & \quad \text{cov}(e_t, e_s) = 0 \text{ for } s \text{ not equal to } t.
\end{align*}
\]

Here, \( e_t \) is called white noise. Clearly a white noise process is stationary, and we will assume that the \( e_t \) are identically and independently distributed. Moreover, if reference to a distribution is necessary, we will assume that they are normally distributed. In other words, they are a Gaussian white noise process.

Four Major Stochastic Processes:

1) Auto Regressive (AR)
2) Moving Average (MA)
3) Auto Regressive Moving Average (ARMA)
4) Auto Regressive Integrated Moving Average (ARIMA)

**Auto Regressive Process:**

Simple representation of time series is Auto Regressive process. In statistics, an autoregressive (AR) model is a representation of a type of random process; as such, it describes certain time-varying processes in economics, etc. The autoregressive model specifies that the output variable depends linearly on its own previous values.

Usually a time series \( L1, x2, \ldots, xt \) generation process will be unknown and even if the process is assumed to be stationary, it can have a more complicated structure than a simple autoregressive (AR) process. AR model is

\[ X_t = \rho x_{t-1} + e_t \]
Here, \( E(x_t) = 0 \) for all \( t \).

\( \text{Var}(x_t) < \infty \)

\( \text{Cov}(x_t, x_{t+k}) = y_k \) for all \( t \) and \( k \).

Consequently, the covariance between two elements \( X_t \) and \( X_{t+k} \) of a time series is

\( \text{Cov}(x_t, x_{t+k}) = \rho^k \sigma^2 x \)

This is called autocovariance because it measures the linear dependence between the members of a single time series.

The covariance of \( X_t \) and \( X_{t+k} \) does not depend on the time point \( T \), but only on the distance the two random variables are apart in time, i.e. on \( K \) since \( \sigma^2 x \) is time invariant. The autocovariance are normalised by dividing each \( y_k \) by the variance \( Y_0 \), of the process to obtain the autocorrelation function.

\( \rho_k = \frac{y_k}{y_0} \) where \( y_0 = \sigma^2 x \)

Partial autocorrelation: One way to identify the order of an adequate AR process for a set of data is to estimate process of increasing order \( K \) and test the significance of \( \Theta_k \). This coefficient is called the \( K \)th partial autocorrelation coefficient and will be denoted by \( \hat{\Theta}_k \), since it is the \( K \)th coefficient of an AR process of order \( K \). It can be shown that for large sample size of the order of the AR is in fact \( q \), the estimated partial autocorrelations \( \hat{\Theta}_k \) are approximately normally distributed with mean 0 and variance \( 1/T \) for \( K>q \), where \( T \) is sample size. Consequently, the significance of the \( \Theta_k \) approximately 95% confidence intervals,

\( (\hat{\Theta}_k -2/\sqrt{T}, \hat{\Theta}_k +2/\sqrt{T}) \)

In this AR process the use of unnecessary many parameters to present the process is inefficient; the question arises whether a more parsimonious representation of this process can be found. Therefore, we will represent an alternative class of stationary stochastic process in the next section.

The estimation of AR models is straightforward. We can estimate them by ordinary least square by minimizing \( \sum e_t^2 \).
Moving Average Process

In the previous section we encountered a process that cannot be represented well by a low order AR process. Consider an infinite AR form:

\[ X_t = -\alpha x_{t-1} - \alpha^2 x_{t-1} - \ldots + e_t \]

Here using lag operator, again \( e_t \) is white noise and \( |\alpha| < 1 \), \( X_t = (1 - \alpha l) e_t \)

A process like that, where \( x_t \) is a weighted sum of members of the white noise series, is called a moving average (MA), since the weighted sum consist only of the member of the white noise series associated with the current and more recent time point. This equation solved by number of parameters and infinite series. This process has a disadvantage which is degrees of freedom is very low. MA and AR both are same process because of stationarity. Here if the condition of the stationarity is satisfied then the data generating process is called invertible.

Given a sample we determine an adequate MA order of the generating process.

The mean is \( E(x_t) = 0 \)

The autocovariances of the MA (p) is

\[ \gamma_k = \sigma^2 \sum_{i=0}^{p-k} \alpha_i \alpha_{i+k} \text{ for } k = 0, 1, 2, \ldots, p \]

\[ = 0 \quad \text{for } k > p \]

Consequently the autocorrelations are

\[ \rho_k = \frac{\sum_{i=0}^{p-k} \alpha_i \alpha_{i+k}}{\sum_{i=0}^{p} \alpha^2} \text{ for } k = 0, 1, \ldots, p \]

\[ = 0 \quad \text{for } k > p \]

To determine whether a particular \( \rho_k \) is nonzero, we can use the available data to compute an estimate of this autocorrelation coefficient and then set up a significance test.

A commonly used estimate for \( \rho_k \) is

\[ r_{k} = \frac{c_k}{c_0} \]

Another possible estimate for \( \rho_k \) is

\[ \hat{r}_k = \frac{\bar{c}}{\bar{c}} \]
The significance of the autocorrelation is often tested by checking whether the $r_k$ or $r_k$ are inside a range $\pm 2/T^{0.5}$. For large $T$ independent of sample mean the $r_k$ and $r_k$ are approximately normally distributed with mean zero and variance $1/T$. For large $T$ if zero does not fall within approximately 95% confidence interval

$$(r_k -2/T^{0.5}, r_k+2/T^{0.5})$$

The null hypothesis $\rho_k = 0$ must be rejected at the 5% level.

**Auto Regressive Moving Average and Auto Regressive Integrated Moving Average**

In ARMA we consider AR and MA process both. If the autocorrelations $\rho_k$ have a cut off point, that is, if they are zero for all $k$ greater than some small number and the partial autocorrelations $\theta_{kk}$ taper of for growing $k$, an MA representation is suggested. AR process is suggested when the autocorrelation taper off and the partial autocorrelations have a cut off point.

The Autoregressive Moving Average Process of order $(q,p)$ is

$$(1-\theta_1L-\theta_2L^2-...-\theta_qL^q)(x_t = (1+\alpha_1L+\alpha_2L^2+...+\alpha_pL^p)e_t$$

One possibility to determine these orders is to use the estimated autocorrelations and partial autocorrelation. In this case in which the autocorrelations die out slowly, the considered process is likely to be nonstationary. Suppose that we start a process

$$Y_t = y_{t-1} + x_t \quad (a)$$

Where $x_t$ is a non-stationary time series with mean $\mu \neq 0$, at time $t=0$, with $y_0=0$ and hence, $E(Y_t) = E(L1+x_2 +..............+x_t) = t\mu$. Thus the mean follows a linear trend and the series is not stationary because stationary requires a constant mean.

To remove the trend we can simply difference $y_t$ and consider

$$X_t = y_t - y_{t-1} = (1-L)y_t \quad (b)$$

It is stationary. Because it appears that differencing is a useful tool to convert nonstationary real-life processes to stationary processes. In equation (a) $x_t$ is an ARMA$(q,p)$ process, then $y_t$ is called an Autoregressive Integrated Moving Average Process which is denoted by ARIMA$(q,1,p)$. If $x_t = (1-L)^d y_t$ is an ARMA$(q,p)$, then $y_t$ is an ARIMA$(q,d,p)$ process, where $d$ is a positive integer. A time series that is stationary after $d$ times differencing is sometimes said to be homogeneous nonstationary of degree $d$. 
BOX JENKINS APPROACH:

1. Differencing the series to achieve stationarity
2. Identify model to be tentatively entertained
3. Estimate the parameters of the tentative model
4. Diagnostic checking. Is the model adequate?
   - No
   - Yes
     - Use the model for forecasting and control
The method is partitioned into three stages:

1. **IDENTIFICATION:**

Identification of the most appropriate model is the most important part of the process. The first step is to determine if the variable is stationary, this can be done with the correlogram. If it is not stationary it needs to be first-differenced. (it may need to be differenced again to induce stationarity). The class of ARMA models is quite large, and in practice we must decide which of these models is most appropriate for the data at hand \(L_1, x_2 \ldots \ldots, x_n\). The chief tools in identification are the autocorrelation function, the partial autocorrelation function and the resulting correlograms.

The next stage is to determine the \(p\) and \(q\) in the ARIMA \((p, I, q)\) model (the \(I\) refers to how many times the data needs to be differenced to produce a stationary series). To determine the appropriate lag structure in the AR part of the model, the PACF or Partial correlogram is used, where the number of non-zero points of the PACF determine where the AR lags need to be included. To determine the MA lag structure, the ACF or correlogram is used, again the non-zero points suggest where the lags should be included. We first describe the correlogram, since it is conceptually the simplest. The theoretical correlogram is a plot of the theoretical autocorrelations

\[
\rho = \text{corr}(x_t, x_{t-k}) \text{ against } k
\]

1). For AR\( (q)\), the partial autocorrelation \(\theta_{kk}\) will be zero for \(k>q\) and autocorrelations taper off. Thus, for \(k\) large (say \(k \geq p\)), the correlogram would be expected to decline steadily. \(\theta_{kk}\) is called the partial correlation between \(x_t\) and \(x_{t-k}\).

A cutoff point of the partial autocorrelation function may be determined by comparing the estimates with \(\pm 2/\sqrt{0.5T}\), since \(1/\sqrt{0.5T}\) is the approximate standard deviation of the estimators \(\hat{\theta}_{kk}\) for \(k>q\).

2). If the series is MA\((p)\) its theoretical correlogram would "cut off" (i.e., take the value zero) for \(k>p\). Thus, we would expect that the sample correlogram would have a similar (though not identical) shape to the theoretical correlogram, and would therefore stay reasonably close to zero for \(k > p\). Reversing this reasoning, we get the rule; if the correlogram seems to cut off for \(k > q\), then the appropriate model is MA\((p)\).

We have already seen some evidence of this: The correlogram for an MA model and the partial correlogram for an AR model both cut off. A still unanswered question is how we can identify a mixed ARMA model. In this case, it can be shown that the correlogram and partial correlogram both die down (but do not cut off). Thus, if both diagrams die down, we can conclude that the
appropriate model is ARMA. Unfortunately, though, the diagrams do not in this case help us to
decide on the order \((p, q)\) of the mixed model.

2. **ESTIMATION:**

This part of the Box-Jenkins methodology is the most straightforward one. Having identified the
appropriate \(p\) and \(q\) values the next stage is to estimate the parameters of the autoregressive and
moving average terms. Some this calculation is done by simple least squares but sometimes it
can be done by some nonlinear estimation methods. The parameters of pure AR processes can be
estimated by using regression methods. It can be estimated by ordinary least square method. If
MA and ARMA are involved the minimization of the sum of squared errors or the maximization
of the likelihood function require nonlinear optimization methods.

3. **DIAGNOSTIC CHECKING:**

Once a model has been identified and estimated, it is usually taken to the the true model and
forecasts can be obtained accordingly. It is virtually certain that the estimated model is not the
true model. To protect against disastrous forecasting errors, the least we can do is to c heck that
the fitted model is a satisfactory one. This is done by the use of diagnostic checks.

In case of Diagnostic checking there are two possibilities:

1) **Over fitting the model:** Let we have specified ARMA \((q,p)\) then estimate either ARMA
\((q+1,p)\) or ARMA \((q,p+1)\) or ARMA \((q,+1p+1)\) and test the significance of extra
parameters.

2) **Residual Analysis:** Compute residual sample autocorrelation \(r_i\) and compute
Portmanteau Test Statistics.

A common test is the Box-Pierce test which is based on the Box-Pierce Q statistics

\[
Q = n \sum_{k=1}^{h} r_k^2
\]

This test was originally developed by Box and Pierce for testing the residuals from a forecast
model. Any good forecast model should have forecast errors which follow a white noise model.
If the series is white noise then, the Q statistic has a chi-square distribution with \(k-q-p\) degrees of
freedom. If it does not follow chi- square then we need overfitting model. Then we again compute Portmanteau Test Statistics.
UNIT ROOT TEST:

The main emphasis was on transforming the data to achieve stationarity and then estimating ARMA models. The differencing operation used to achieve a stationary involves a loss of potential information about long-run movements. The Box- Jenkins method of differencing the time series after a visual inspection of the correlogram has been formalized in the tests for unit roots. The literature on unit roots studies nonstationary which is stationary in first difference.

Consider the model,

\[ y_t = \alpha y_{t-1} + \varepsilon_t \]

Where, \( \varepsilon_t \) is white noise. In the random walk case \( (\alpha=1) \) it is well known that the OLS estimation of this equation produces an estimate of \( \alpha \) that is biased toward zero. However, the OLS estimate is also biased toward zero when \( \alpha \) is less than but near to zero. Evans and Savin provide Monte Carlo evidence on the bias and the other aspects of the distributions.

To discuss the Dicky-Fuller tests, consider the model

\[ y_t = \beta_0 + \beta_1 t + u_t \]
\[ u_t = \alpha u_{t-1} + \varepsilon_t \]

Where, \( \varepsilon_t \) is a covariance stationary process with zero mean. The reduced form for this model is

\[ y_t = \gamma + \delta t + \alpha y_{t-1} + \varepsilon_t \]

Where, \( \gamma = \beta_0 (1-\alpha) + \beta_1 \alpha \). This equation is said to have a unit root if \( \alpha=1 \) (in which case \( \delta =0 \)).

DICKEY – FULLER TEST:

The Dickey-Fuller tests are based on testing the hypothesis \( \alpha=1 \) in equation 1 under the assumption that \( \varepsilon_t \) are white noise errors. There are three test statistics

\[ K(1) = T (\hat{\alpha} - 1) \]
\[ t(1) = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})} F (0,1) \]

where \( \hat{\alpha} \) is the OLS estimate of \( \alpha \) in equation 1, \( SE(\hat{\alpha}) \) is the standard error of \( \hat{\alpha} \) and \( F(0,1) \) is the usual F- statistics for testing the joint hypothesis \( \delta=0 \) and \( \alpha=1 \) in equation 1. These statistics do not have the standard normal t and F distribution. The critical values for \( K(1) \) and \( t(1) \) are
tabulated for $\delta=0$ in Fuller and the critical values for the $F(0,1)$ statistics are tabulated in Dickey and Fuller (1981).

**THE SERIAL CORRELATION PROBLEM:**

Dickey, Fuller and others developed modifications for Dickey fuller tests when $e_t$ is not a white noise. It is called augmented Dickey Fuller test.

Augmented Dickey Fuller Test:

$$Y_t = \gamma + \delta t + \alpha y_{t-1} + \sum_{j=1}^{k} \theta_j \Delta y_{t-j} + \varepsilon_t$$

Where, $\Delta y_{t-j}$ - take into account ARMA effect and $T = $ sample size.

**STATIONARY PROCESS:**

Suppose we have some trend in the series. There are two major ways of detrending the series.

1) Trend stationary process
2) Difference stationary process

Trend stationary process:

$$Y_t = f(t) + u_t \quad f(t) = \alpha + \beta t \text{ (linear trend)}$$

$$Y_t = \alpha + \beta t + u_t .$$

So, there is trend in the equation. We remove this trend. Apply OLS to the above equation

$$\hat{Y} = \hat{\alpha} + \hat{\beta} t. \quad \hat{Y} \text{ is estimated trend path.}$$

Compute $\hat{u}_t = Y_t - \hat{Y}_t.$ In this equation there is no trend.

$\hat{u}_t$ is detrended series. It satisfies a) $\hat{u}_t = 0$ and $\sum t \hat{u}$

Nelson and Ploiser called this model Trend Stationarity Process.

$$Y_t = \alpha + \beta t + u_t$$

$\text{E}(Y_t) = \alpha + \beta t.$

$\text{Var}(Y_t) = \sigma_u^2$
In case of TSP, mean depends on trend but variance does not depend on trend.

**Difference stationary process**

Here differencing is needed to obtain stationary.

\[ Y_t = \alpha + \beta t + u_t. \]

\[ \Delta Y_t = \beta t + u_t - u_{t-1} \quad \text{so,} \quad Y_t - Y_{t-1} = \beta + \varepsilon_t. \]

Such a process is called Random Walk with drift. If the random walk model predicts that the value at time "t" will equal the last period's value plus a constant, or drift (\(\beta\)), and a white noise term (\(\varepsilon_t\)), then the process is random walk with a drift. It also does not revert to a long-run mean and has variance dependent on time.

\[ \Delta^2 Y_t = u_t - 2u_{t-1} + u_{t-2}. \]

Such a process is called Random Walk without drift. Random walk predicts that the value at time "t" will be equal to the last period value plus a stochastic (non-systematic) component that is a white noise, which means \(\varepsilon_t\) is independent and identically distributed with mean "0" and variance "\(\sigma^2\)". Random walk can also be named a process integrated of some order, a process with a unit root or a process with a stochastic trend. It is a non mean reverting process that can move away from the mean either in a positive or negative direction. Another characteristic of a random walk is that the variance evolves over time and goes to infinity as time goes to infinity; therefore, a random walk cannot be predicted.

\[ Y_t - Y_{t-1} = \beta + \varepsilon_t \]

\[ E(Y_t - Y_{t-1}) = \beta \]

\[ \text{Var} (Y_t) = t\sigma^2. \] In case of DSP, mean does not depend on trend but variance depends on trend.

Test for TSP and DSP

\[ Y_t = \alpha + \rho Y_{t-1} + \beta t + \varepsilon_t \]

If \(\rho = 1\) and \(\beta = 0\), the equation have no trend which implies DSP. If \(\rho < 1\) it implies TSP.

**COINTEGRATION:**

An important issue in econometrics is the need to integrate short run dynamics with long run equilibrium. The theory of co integration explains how to study the interrelationship between the
long term trends in the variables, trends that are differenced away in the Box-Jenkins methods. This procedure, however, throws away potential valuable information about long run relationships about which economic theories have a lot to say.

A time series $y_t$ is said to be integrated of order 1 or I(1) if $\Delta y_t$ is a stationary time series. A stationary time series is said to be I(0). A random walk is a special case of an I(1) series, because if $y_t$ is a random walk, $\Delta y_t$ is a random series or white noise. If $y_t$ I(1), and $\sim$I(0), then their sum $Z_t = y_t + u_t \sim I(1)$.

Suppose, $y_t-\beta x_t$ is I(0). This is denoted by saying $y_t$ and $x_t$ are CI (1,1). This means that the regression equation $y_t=\beta x_t + u_t$.

This makes sense because $y_t$ and $x_t$ do not drift too far apart from each other over time. Thus there is a long run equilibrium relationship between them. If $y_t$ and $x_t$ are not cointegrated, that is $y_t-\beta x_t = u_t$. It is also I(1), they can drift apart from each other mare and mare as time goes on. Thus, there is no long run equilibrium relationship between them.

**Definition of Cointegration:** Suppose that $\sim$I(1), $\sim$I(1). Then $y_t$ and $x_t$ are said to be cointegrated if there exists a $\beta$ such that $y_t-\beta x_t$ is I(0). Cointegration relation is a long run relationship.

**TEST FOR COINTEGREATION:**

1) Apply unit root on $y_t$ to see whether $y_t$ is I(1).
2) Apply unit root on $x_t$ to see whether $x_t$ is I(1).
3) Apply OLS to obtain $\hat{u_t}$.
4) Apply unit root on $\hat{u_t}$ to see whether it is I(0).

Two step method of CI, which is

a) If $y$ and $x$ are CI then we get long run relationship.

b) If $y$ and $x$ is distributed then we get fluctuation computed by error correction method.

Error correction is done when a stable longrun trend exists and some kind of fluctuation is generated around it. And if stable longrun trend does not exists there error correction is not needed.

**Bewly and Wicknes and Brunch** showed that shortrun and longrun relationship can be estimated simultaneously. Now equation can be endogenous which is
$Y_t = \beta x_t + \Delta y_t - \Delta x_t - \nu_t / \lambda$

This equation can be used for simultaneous estimators, but OLS is inappropriate. Use IV method. Use error correction method. Cointegration also support error correction method.

TESTING GOODNESS OF FIT FOR TIME SERIES MODEL

1) Akaike Information Criteria (AIC)
   2) Schwarz Bayesian Criteria (SBC)

The equations are:

$$AIC (P) = n \log \tilde{\sigma}_p^2 + 2p$$

$$BIC (P) = n \log \tilde{\sigma}_p^2 + p \log n$$

Where, $n$= sample size, $p$ = total no of parameter, $\tilde{\sigma}_p^2$=RSS/n-p, RSS= Residual Sum Of Square Error. The model is chosen in the manner which is best fitted for which AIC or BIC minimum.

4. DATA

We have taken data on four variables - food grains production in India (F1), major commercial crop production in India (C1), Total Agricultural Production (C1+F1=T1) & institutional credit to the agricultural sector (L1) for the years 1970-2008. Commercial crops include oil seeds, cotton, raw jute & mesta, sugarcane, tobacco. The data on crop production is given in million tons while the data on agricultural credit is given in Rs. Crores. I look into the concepts of stationarity and cointegration properties of our chosen datasets. The datasets are taken from the official website of the Reserve Bank of India. The data have been collected from the Reserve Bank of India database- time series publications, Handbook of Statistics on the Indian economy. The variables referred in the following empirical results denote the following:

L1: Total loans given to the agricultural sector by the institutional credit sources. It includes both loans issued and loans outstanding for years 1970-2008

F1: Total foodgrain production for years 1970-2008

C1: Total commercial crop production for years 1970-2008

T1: Total agricultural production for years 1970-2008 which includes both commercial and foodgrain production.
4. ECONOMETRIC ANALYSIS

CORRELOGRAM:

Date: 10/29/13  Time: 11:36  
Sample: 1970 2008  
Included observations: 38

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
<td>0.905</td>
<td>0.905</td>
<td>34.459</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2</td>
<td>0.799</td>
<td>-0.111</td>
<td>62.031</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3</td>
<td>0.636</td>
<td>-0.094</td>
<td>82.915</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>4</td>
<td>0.576</td>
<td>-0.047</td>
<td>98.052</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>5</td>
<td>0.459</td>
<td>-0.106</td>
<td>107.97</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>6</td>
<td>0.382</td>
<td>0.029</td>
<td>114.30</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>7</td>
<td>0.281</td>
<td>-0.086</td>
<td>117.70</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>8</td>
<td>0.178</td>
<td>0.020</td>
<td>119.34</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>9</td>
<td>0.115</td>
<td>0.034</td>
<td>120.05</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>10</td>
<td>0.065</td>
<td>-0.018</td>
<td>120.28</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>11</td>
<td>0.024</td>
<td>0.001</td>
<td>120.31</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>12</td>
<td>-0.012</td>
<td>-0.045</td>
<td>120.32</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>13</td>
<td>-0.044</td>
<td>-0.030</td>
<td>120.44</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>14</td>
<td>-0.075</td>
<td>-0.035</td>
<td>120.80</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>15</td>
<td>-0.096</td>
<td>0.001</td>
<td>121.42</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>16</td>
<td>-0.117</td>
<td>-0.029</td>
<td>122.38</td>
</tr>
</tbody>
</table>

Fig 1: The above figure shows the correlogram of the dependent variable of our model i.e total loans (L1) to the agricultural sector from the institutional sources. The correlogram is done to get a basic idea of which time series process the variable is following. Here we find that total loans follow an AR(1) process
Fig 2: The above figure shows the correlogram of the 1st independent variable i.e. foodgrains production (F1). Here we find that total foodgrains production follow an AR(1) process.
Fig 3: The above figure shows the correlogram of the 2nd independent variable i.e commercial crop production (C1). It follows AR(1).
Fig 4: The above figure shows the correlogram of the 3rd independent variable i.e total agricultural production(T1) which includes both the foodgrains as well as commercial crop production.
L1 correlogram (estimation)

Dependent Variable: SER05  
Method: Least Squares  
Date: 10/29/13  Time: 11:51  
Sample (adjusted): 1971 2008  
Included observations: 38 after adjustments  
Convergence achieved after 3 iterations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.093800</td>
<td>0.014066</td>
<td>77.76252</td>
<td>0.0000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.990353</td>
<td>Mean dependent var</td>
<td>1761.256</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.990353</td>
<td>S.D. dependent var</td>
<td>2331.033</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>228.9543</td>
<td>Akaike info criterion</td>
<td>13.73089</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>1939543</td>
<td>Schwarz criterion</td>
<td>13.77398</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-259.8868</td>
<td>Hannan-Quinn criter.</td>
<td>13.74622</td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.442789</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inverted AR Roots  
1.09  
Estimated AR process is nonstationary

Fig 5: By estimating the variable ‘total loans’ (T1) we find that the estimated AR process in non-stationary. So, we need to difference it till the AR process is stationary.
ESTIMATION:

Dependent Variable: DSER05  
Method: Least Squares  
Date: 10/29/13  Time: 16:13  
Sample (adjusted): 1973 2008  
Included observations: 36 after adjustments  
Convergence achieved after 13 iterations  
MA Backcast: 1972

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.330427</td>
<td>0.222829</td>
<td>1.482870</td>
<td>0.1476</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.603575</td>
<td>0.188617</td>
<td>3.199996</td>
<td>0.0030</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.145938</td>
<td>0.288609</td>
<td>-0.505659</td>
<td>0.6165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.553278</td>
<td>Mean dependent var</td>
<td>202.4517</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.526203</td>
<td>S.D. dependent var</td>
<td>282.5907</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>194.5153</td>
<td>Akaike info criteron</td>
<td>13.45855</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>1248595.</td>
<td>Schwarz criterion</td>
<td>13.59051</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-239.2540</td>
<td>Hannan-Quinn criter.</td>
<td>13.50461</td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.953810</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inverted AR Roots</td>
<td>.96</td>
<td>-.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inverted MA Roots</td>
<td>.15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 7: Here we find that L1 follows ARMA(2,1). Correlogram for DF1(i.e first difference of D1)
Date: 10/29/13  Time: 12:05
Sample: 1970 2008
Included observations: 38

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-0.633</td>
<td>-0.633</td>
<td>16.480</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.300</td>
<td>-0.170</td>
<td>20.271</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>-0.104</td>
<td>0.018</td>
<td>20.742</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>-0.145</td>
<td>-0.282</td>
<td>21.687</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.215</td>
<td>-0.064</td>
<td>23.813</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>-0.201</td>
<td>-0.064</td>
<td>25.740</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>0.116</td>
<td>-0.113</td>
<td>26.398</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.026</td>
<td>0.046</td>
<td>26.431</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>-0.074</td>
<td>0.035</td>
<td>26.719</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>-0.020</td>
<td>-0.217</td>
<td>26.742</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>0.125</td>
<td>0.090</td>
<td>27.618</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>-0.221</td>
<td>-0.107</td>
<td>30.481</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>0.224</td>
<td>-0.087</td>
<td>33.535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>-0.227</td>
<td>-0.151</td>
<td>36.809</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>0.190</td>
<td>0.004</td>
<td>39.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>-0.018</td>
<td>0.059</td>
<td>39.217</td>
</tr>
</tbody>
</table>
Estimation of DF1 (1st difference of F1):

Dependent Variable: DSER06  
Method: Least Squares  
Date: 10/29/13  Time: 12:13  
Sample (adjusted): 1973 2008  
Included observations: 36 after adjustments  
Convergence achieved after 15 iterations  
MA Backcast: 1972

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.351067</td>
<td>0.133018</td>
<td>2.639249</td>
<td>0.0126</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.632871</td>
<td>0.135023</td>
<td>4.687129</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.968855</td>
<td>0.021640</td>
<td>-44.77095</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.408499  Mean dependent var 3.817778  
Adjusted R-squared 0.372650  S.D. dependent var 14.85163  
S.E. of regression 11.76330  Akaike info criterion 7.847501  
Sum squared resid 4566.382  Schwarz criterion 7.979461  
Log likelihood -138.2550  Hannan-Quinn criter. 7.893559  
Durbin-Watson stat 2.261596

Inverted AR Roots .99  -.64  
Inverted MA Roots .97

Fig 8: DF1 follows ARMA (2,1).
Correlogram of DF2 (1st difference of F2)

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0.575</td>
<td>0.575</td>
<td>13.595</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 0.720</td>
<td>0.581</td>
<td>35.473</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 0.544</td>
<td>0.081</td>
<td>48.313</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 0.427</td>
<td>-0.293</td>
<td>56.478</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 0.385</td>
<td>-0.084</td>
<td>63.303</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 0.251</td>
<td>-0.000</td>
<td>66.307</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 0.163</td>
<td>-0.112</td>
<td>67.611</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 0.037</td>
<td>-0.214</td>
<td>67.681</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 0.077</td>
<td>0.238</td>
<td>67.991</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 -0.024</td>
<td>0.194</td>
<td>68.022</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 -0.012</td>
<td>-0.110</td>
<td>68.030</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 -0.060</td>
<td>-0.198</td>
<td>68.237</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 -0.097</td>
<td>-0.015</td>
<td>68.809</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 -0.103</td>
<td>0.023</td>
<td>69.476</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 -0.120</td>
<td>-0.083</td>
<td>70.427</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 -0.132</td>
<td>-0.056</td>
<td>71.625</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Estimation of DC1 (1st difference of C1)

Dependent Variable: DSER07  
Method: Least Squares  
Date: 10/29/13  Time: 12:19  
Sample (adjusted): 1973 2008  
Included observations: 36 after adjustments  
Convergence achieved after 14 iterations  
MA Backcast: 1972

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.328761</td>
<td>0.231903</td>
<td>1.417662</td>
<td>0.1657</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.596473</td>
<td>0.194988</td>
<td>3.059048</td>
<td>0.0044</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.108949</td>
<td>0.301679</td>
<td>-0.361142</td>
<td>0.7203</td>
</tr>
</tbody>
</table>

R-squared          0.556354  
Adjusted R-squared 0.529467  
S.E. of regression 488.0710  
Akaike info criterion 15.29845  
Sum squared resid  7861037.  
Schwarz criterion   15.43041  
Log likelihood     -272.3722  
Hannan-Quinn criter. 15.34451  
Durbin-Watson stat  1.930046

Inverted AR Roots .95  
Inverted MA Roots .11

Fig 10: DC1 follows ARMA(2,1)
Estimation of DT1 (1st difference of T1)

Dependent Variable: DSER08
Method: Least Squares
Date: 10/29/13  Time: 12:26
Sample (adjusted): 1973 2008
Included observations: 36 after adjustments
Convergence achieved after 13 iterations
MA Backcast: 1972

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.340783</td>
<td>0.236766</td>
<td>1.439320</td>
<td>0.1595</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.586723</td>
<td>0.199838</td>
<td>2.935992</td>
<td>0.0060</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.113225</td>
<td>0.305431</td>
<td>-0.370705</td>
<td>0.7132</td>
</tr>
</tbody>
</table>

R-squared     0.559269  Mean dependent var 523.9303
Adjusted R-squared 0.532558  S.D. dependent var 707.4377
S.E. of regression 483.6733    Akaike info criterion 15.28035
Sum squared resid 7720015.  Schwarz criterion 15.41231
Log likelihood -272.0463  Hannan-Quinn criter. 15.32641
Durbin-Watson stat 1.929461

Inverted AR Roots .96  -.61
Inverted MA Roots .11

Fig 11: DT1 follows ARMA (2,1)
Correlogram estimation of C1

Dependent Variable: SER07
Method: Least Squares
Date: 10/29/13 Time: 12:30
Sample (adjusted): 1971 2008
Included observations: 38 after adjustments
Convergence achieved after 3 iterations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.090011</td>
<td>0.013517</td>
<td>80.64277</td>
<td>0.0000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.990731</td>
<td>Mean dependent var</td>
<td>4757.886</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.990731</td>
<td>S.D. dependent var</td>
<td>6034.860</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>580.9893</td>
<td>Akaike info criterion</td>
<td>15.59331</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>12489317</td>
<td>Schwarz criterion</td>
<td>15.63640</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-295.2728</td>
<td>Hannan-Quinn criter.</td>
<td>15.60864</td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.341383</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inverted AR Roots 1.09
Estimated AR process is non-stationary

Fig 12: the above figure shows that the estimated AR process of C1 is non-stationary. So, we need to difference it to obtain stationarity.
Correlogram Estimation of T1:

Dependent Variable: SER08
Method: Least Squares
Date: 10/29/13 Time: 12:31
Sample (adjusted): 1971 2008
Included observations: 38 after adjustments
Convergence achieved after 3 iterations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>1.089265</td>
<td>0.013103</td>
<td>83.12902</td>
<td>0.0000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.991073</td>
<td>Mean dependent var</td>
<td>4922.252</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.991073</td>
<td>S.D. dependent var</td>
<td>6065.994</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>573.1386</td>
<td>Akaike info criterion</td>
<td>15.56610</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>12154052</td>
<td>Schwarz criterion</td>
<td>15.60919</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-294.7558</td>
<td>Hannan-Quinn criter.</td>
<td>15.58143</td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.325590</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inverted AR Roots 1.09
Estimated AR process is nonstationary

Fig 13: the above figure shows that the estimated AR process of T1 is non-stationary. So, we need to difference it in order to obtain stationarity.

The stationarity of L1 is obtained at the fourth differencing i.e at DDDDL1.

Correlogram Analysis for Stationarity

The stationarity of the variables has been examined through the study of their Auto-correlation functions (ACFs) & Partial Auto-correlation Functions (PACFs). Figures 1-4 represent the ACF & PACFs of the variables concerned. It is observed from these functions that the series for L1,C1, F1 & T1 are non-stationary.

Test of Unit Roots

In case of time series analysis, unit root tests are important since these tests detect the stationarity and non-stationarity of the time series data used for the study. Regression run on non-stationary time series produces spurious relations. To avoid this, it becomes necessary to perform a unit root test on the variables. The Augmented Dickey-Fuller (ADF) test is widely used for performing unit root test.
Now, I perform the unit root test to detect whether the datasets are stationary or non-stationary.

**Augmented Dickey-Fuller Tests for unit roots**

**UNIT ROOT TEST:**

**Unit Root Test of DDDDL1:**

| Null Hypothesis: DDDDSER05 has a unit root |  |
| Exogenous: Constant, Linear Trend |  |
| Lag Length: 5 (Automatic - based on AIC, maxlag=8) |  |
| | **t-Statistic** | **Prob.** |
| Augmented Dickey-Fuller test statistic | -10.48618 | 0.0000 |
| Test critical values: |  |
| 1% level | -4.309824 |
| 5% level | -3.574244 |
| 10% level | -3.221728 |


**Augmented Dickey-Fuller Test Equation**

| Dependent Variable: D(DDDDSER05) |  |
| Method: Least Squares |  |
| Date: 10/29/13 Time: 12:42 |  |
| Sample (adjusted): 1980 2008 |  |
| Included observations: 29 after adjustments |  |

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>D(DDDDSER05(-1))</td>
<td>-16.73558</td>
<td>1.595965</td>
<td>-10.48618</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER05(-1))</td>
<td>13.44409</td>
<td>1.528904</td>
<td>8.793287</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER05(-2))</td>
<td>10.49826</td>
<td>1.326938</td>
<td>7.911641</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER05(-3))</td>
<td>7.496121</td>
<td>1.000930</td>
<td>7.489159</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER05(-4))</td>
<td>4.573456</td>
<td>0.626249</td>
<td>7.302938</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER05(-5))</td>
<td>1.839464</td>
<td>0.245426</td>
<td>7.494972</td>
<td>0.0000</td>
</tr>
<tr>
<td>C</td>
<td>93.67744</td>
<td>91.71968</td>
<td>1.021345</td>
<td>0.3187</td>
</tr>
<tr>
<td>@TREND(&quot;1970&quot;)</td>
<td>-5.042592</td>
<td>3.666167</td>
<td>-1.375440</td>
<td>0.1835</td>
</tr>
</tbody>
</table>

| R-squared | 0.995355 | Mean dependent var | 21.78172 |
| Adjusted R-squared | 0.993806 | S.D. dependent var | 2023.585 |
| S.E. of regression | 159.2572 | Akaike info criterion | 13.20787 |
| Sum squared resid | 532620.2 | Schwarz criterion | 13.58505 |
| Log likelihood | -183.5141 | Hannan-Quinn criter. | 13.32600 |
| F-statistic | 642.8096 | Durbin-Watson stat | 1.661126 |
| Prob(F-statistic) | 0.000000 |  |

H0: DDDDL1 has a unit root. H1: DDDDL1 does not have a unit root.
The above figure shows the null hypothesis is rejected so it implies that DDDDL1 does not have a unit root.

**Unit Root test of DF1:**

Null Hypothesis: DSER06 has a unit root  
Exogenous: Constant, Linear Trend  
Lag Length: 0 (Automatic - based on AIC, maxlag=9)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>-12.43035</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Test critical values:  
1% level: -4.226815  
5% level: -3.536601  
10% level: -3.200320


Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(DSER06)  
Method: Least Squares  
Date: 10/29/13 Time: 12:45  
Sample (adjusted): 1972 2008  
Included observations: 37 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSER06(-1)</td>
<td>-1.637421</td>
<td>0.131728</td>
<td>-12.43035</td>
<td>0.0000</td>
</tr>
<tr>
<td>C</td>
<td>3.637699</td>
<td>4.088125</td>
<td>0.889821</td>
<td>0.3798</td>
</tr>
<tr>
<td>@TREND(&quot;1970&quot;)</td>
<td>0.089243</td>
<td>0.180313</td>
<td>0.544849</td>
<td>0.5894</td>
</tr>
</tbody>
</table>

R-squared: 0.819665  
Adjusted R-squared: 0.809057  
S.E. of regression: 11.69207  
Sum squared resid: 4647.953  
Log likelihood: -141.9161  
F-statistic: 77.26877  
Prob(F-statistic): 0.000000  
Mean dependent var: 0.187568  
S.D. dependent var: 26.75710  
Akaike info criterion: 7.833303  
Schwarz criterion: 7.963918  
Hannan-Quinn crit.: 7.879351  
Durbin-Watson stat: 2.211121
Null Hypothesis: DDDDSER07 has a unit root  
Exogenous: Constant, Linear Trend  
Lag Length: 5 (Automatic - based on AIC, maxlag=8)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-10.13631</td>
</tr>
</tbody>
</table>

Test critical values:  
- 1% level: -4.309824  
- 5% level: -3.574244  
- 10% level: -3.221728


Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(DDDDSER07)  
Method: Least Squares  
Date: 10/29/13  Time: 15:34  
Sample (adjusted): 1980 2008  
Included observations: 29 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDDDSER07(-1)</td>
<td>-16.71926</td>
<td>1.649442</td>
<td>-10.13631</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER07(-1))</td>
<td>13.40123</td>
<td>1.575877</td>
<td>8.503981</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER07(-2))</td>
<td>10.39949</td>
<td>1.359722</td>
<td>7.648250</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER07(-3))</td>
<td>7.334895</td>
<td>1.017232</td>
<td>7.210639</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER07(-4))</td>
<td>4.407022</td>
<td>0.632058</td>
<td>6.972502</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDDSER07(-5))</td>
<td>1.735904</td>
<td>0.247985</td>
<td>7.000040</td>
<td>0.0000</td>
</tr>
<tr>
<td>C</td>
<td>233.1616</td>
<td>239.2379</td>
<td>0.974601</td>
<td>0.3409</td>
</tr>
<tr>
<td>@TREND(&quot;1970&quot;)</td>
<td>-12.48000</td>
<td>9.550032</td>
<td>-1.306802</td>
<td>0.2054</td>
</tr>
</tbody>
</table>

R-squared: 0.994665  
Mean dependent var: 63.76379  
Adjusted R-squared: 0.992287  
S.D. dependent var: 4932.431  
S.E. of regression: 416.0041  
Akaike info criterion: 15.12822  
Sum squared resid: 3634247  
Schwarz criterion: 15.50540  
Log likelihood: -211.3592  
Hannan-Quinn criterion: 15.24635  
F-statistic: 559.3245  
Durbin-Watson stat: 1.680556  
Prob(F-statistic): 0.000000
Unit root test of DDDDT1

Fig 17: DDDDT1 has no unit root because the null hypothesis which states that there is unit root is rejected.
Correlogram of DDDDT1

Date: 10/29/13  Time: 15:39  
Sample: 1970 2008  
Included observations: 35

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-0.853 -0.853</td>
<td>27.702</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.525  -0.742</td>
<td>39.518</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>-0.213 -0.538</td>
<td>40.358</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.027  -0.289</td>
<td>40.308</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.035  -0.157</td>
<td>40.440</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>-0.053 -0.450</td>
<td>40.568</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>0.108  -0.231</td>
<td>41.107</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>-0.178 0.126</td>
<td>42.627</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>0.199  0.176</td>
<td>44.606</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>-0.161 -0.021</td>
<td>45.943</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>0.091  -0.080</td>
<td>46.393</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>-0.028 0.007</td>
<td>46.438</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>-0.006 -0.077</td>
<td>46.439</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>0.016  -0.230</td>
<td>46.455</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>-0.019 -0.109</td>
<td>46.479</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>0.015  -0.001</td>
<td>46.494</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Correlogram of DDDDC1:

Date: 10/29/13  Time: 15:41  
Sample: 1970 2008  
Included observations: 35

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-0.857 -0.857</td>
<td>27.984</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.538  -0.740</td>
<td>39.356</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>-0.232 -0.532</td>
<td>41.539</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.043  -0.292</td>
<td>41.616</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.027  -0.155</td>
<td>41.649</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>-0.056 -0.448</td>
<td>41.790</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>0.117  -0.220</td>
<td>42.418</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>-0.188 0.146</td>
<td>44.105</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>0.207  0.181</td>
<td>46.233</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>-0.166 -0.034</td>
<td>47.660</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>0.097  -0.089</td>
<td>48.168</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>-0.035 0.006</td>
<td>48.237</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>0.002  -0.069</td>
<td>48.237</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>0.008  -0.241</td>
<td>48.241</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>-0.010 -0.109</td>
<td>48.248</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>0.006  0.023</td>
<td>48.250</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Correlogram of DDDDL1

Date: 10/29/13  Time: 15:43
Sample: 1970-2008
Included observations: 35

<table>
<thead>
<tr>
<th>Autocorrelation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.858</td>
<td>-0.858</td>
<td>28.059</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.540</td>
<td>-0.747</td>
<td>39.503</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.234</td>
<td>-0.569</td>
<td>41.712</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.045</td>
<td>-0.330</td>
<td>41.798</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.026</td>
<td>-0.157</td>
<td>41.827</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-0.058</td>
<td>-0.479</td>
<td>41.984</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.121</td>
<td>-0.396</td>
<td>42.664</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-0.187</td>
<td>-0.076</td>
<td>44.336</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.200</td>
<td>0.092</td>
<td>46.323</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.157</td>
<td>-0.046</td>
<td>47.595</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.091</td>
<td>-0.098</td>
<td>48.043</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-0.037</td>
<td>0.043</td>
<td>48.119</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.010</td>
<td>0.067</td>
<td>48.125</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-0.003</td>
<td>-0.163</td>
<td>48.126</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.001</td>
<td>-0.167</td>
<td>48.126</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-0.005</td>
<td>-0.075</td>
<td>48.127</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Estimation of DDDDL1:

Dependent Variable: DDDDSER05
Method: Least Squares
Date: 10/29/13  Time: 15:53
Sample (adjusted): 1975-2008
Included observations: 34 after adjustments
Convergence achieved after 19 iterations
MA Backcast: OFF  (Roots of MA process too large)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>-0.840178</td>
<td>0.100636</td>
<td>-8.366964</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-1.096673</td>
<td>0.082532</td>
<td>-13.287800</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared          0.918746  Mean dependent var  0.333235
Adjusted R-squared 0.916207  S.D. dependent var  968.8208
S.E. of regression  280.4452  Akaike info criterion 14.16766
Sum squared resid   2516783.  Schwarz criterion  14.25744
Log likelihood     -238.8501  Hannan-Quinn crit.  14.19828
Durbin-Watson stat  2.848092

Inverted AR Roots  -.84
Inverted MA Roots  1.10
Estimated MA process is noninvertible
The above table shows that DDDDL1 follows ARMA (1,1)

**Estimation of DDDDC1:**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>-0.836055</td>
<td>0.099054</td>
<td>-8.440403</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-1.135762</td>
<td>0.028966</td>
<td>-39.20980</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- R-squared: 0.924563, Mean dependent var: 6.410882
- Adjusted R-squared: 0.922205, S.D. dependent var: 2364.453
- S.E. of regression: 659.4856, Akaike info criterion: 15.87782
- Sum squared resid: 13917480, Schwarz criterion: 15.96761
- Log likelihood: -267.9229, Hannan-Quinn criter.: 15.90844
- Durbin-Watson stat: 2.922874

Inverted AR Roots: -0.84
Inverted MA Roots: 1.14

Estimated MA process is noninvertible

The above table shows that DDDDC1 follows ARMA(1,1)
Estimation of DDDDT1:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>-0.837229</td>
<td>0.100357</td>
<td>-8.342542</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-1.095394</td>
<td>0.029384</td>
<td>-36.53259</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

The above table shows that DDDDT1 follows ARMA(1,1)

COINTEGRATION:

Cointegration between L1 and F1:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>100.7367</td>
<td>379.1295</td>
<td>0.265705</td>
<td>0.7921</td>
</tr>
<tr>
<td>SER06</td>
<td>-0.639824</td>
<td>2.203399</td>
<td>-0.299227</td>
<td>0.7736</td>
</tr>
</tbody>
</table>

The above table shows that DDDDT1 follows ARMA(1,1)
The above table shows that the cointegration between L1 and F1. There is no cointegration in the long run between L1 and F1.
Cointegration between L1 and C1:

Dependent Variable: DDDSER05
Method: Least Squares
Date: 10/29/13  Time: 16:39
Sample (adjusted): 1973 2008
Included observations: 36 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.340589</td>
<td>4.644444</td>
<td>-0.073018</td>
<td>0.9422</td>
</tr>
<tr>
<td>DDDSER07</td>
<td>0.409030</td>
<td>0.003938</td>
<td>103.8681</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.996858  Mean dependent var -6.509444
Adjusted R-squared 0.996766  S.D. dependent var 492.0931
S.E. of regression 27.90440  Akaike info criterion 9.555124
Sum squared resid 26626.30  Schwarz criterion 9.643097
Log likelihood -169.9922  Hannan-Quinn criterion 9.585829
F-statistic 10738.58  Durbin-Watson stat 3.114259
Prob(F-statistic) 0.000000
Unit root test

Null Hypothesis: RDDSER05 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on AIC, maxlag=9)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test</td>
<td>-8.313952</td>
<td>0.0000</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-4.273277</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.557759</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.212361</td>
<td></td>
</tr>
</tbody>
</table>


Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RDDSER05)
Method: Least Squares
Date: 10/29/13 Time: 16:45
Sample (adjusted): 1977 2008
Included observations: 32 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDDSER05(-1)</td>
<td>-5.637158</td>
<td>0.678036</td>
<td>-8.313952</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(RDDSER05(-1))</td>
<td>3.254854</td>
<td>0.585353</td>
<td>5.560496</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(RDDSER05(-2))</td>
<td>1.745172</td>
<td>0.380592</td>
<td>4.656295</td>
<td>0.0001</td>
</tr>
<tr>
<td>D(RDDSER05(-3))</td>
<td>0.738185</td>
<td>0.161197</td>
<td>4.579393</td>
<td>0.0001</td>
</tr>
<tr>
<td>C</td>
<td>-6.181570</td>
<td>5.646106</td>
<td>-1.094638</td>
<td>0.2836</td>
</tr>
<tr>
<td>@TREND(&quot;1970&quot;)</td>
<td>0.473327</td>
<td>0.236746</td>
<td>1.999303</td>
<td>0.0561</td>
</tr>
</tbody>
</table>

R-squared        0.952536  Mean dependent var  -2.839256
Adjusted R-squared 0.943409  S.D. dependent var  50.67981
S.E. of regression 12.05619  Akaike info criterion 7.984393
Sum squared resid  3779.142  Schwarz criterion  8.259219
Log likelihood    -121.7503  Hannan-Quinn criter. 8.075490
F-statistic       104.3574  Durbin-Watson stat  2.079802
Prob(F-statistic) 0.000000

Fig 25: This figure shows the cointegration between L1 and C1. There is cointegration in the long run between L1 and C1.
Cointegration between L1 and T1:

Dependent Variable: DDDSER05
Method: Least Squares
Date: 10/29/13 Time: 16:46
Sample (adjusted): 1973 2008
Included observations: 36 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.159127</td>
<td>7.210260</td>
<td>-0.022070</td>
<td>0.9825</td>
</tr>
<tr>
<td>DDDSER08</td>
<td>0.417281</td>
<td>0.006224</td>
<td>67.04716</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared      | 0.992493     | Mean dependent var | -6.509444 |
Adjusted R-squared | 0.992273 | S.D. dependent var | 492.0931  |
S.E. of regression | 43.25783 | Akaike info criterion | 10.42619  |
Sum squared resid  | 63622.14  | Schwarz criterion   | 10.51416  |
Log likelihood    | -185.6714   | Hannan-Quinn criter. | 10.45689  |
F-statistic       | 4495.322    | Durbin-Watson stat  | 3.356936  |
Prob(F-statistic) | 0.000000    |                      |           |

Unit root test:

Null Hypothesis: DDDSER05 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 3 (Automatic - based on AIC, maxlag=9)

<table>
<thead>
<tr>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-7.126879</td>
</tr>
</tbody>
</table>

Test critical values:
1% level | -4.273277
5% level | -3.557759
10% level | -3.212361


Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DDDSER05)
Method: Least Squares
Date: 10/29/13 Time: 16:48
Sample (adjusted): 1977 2008
Included observations: 32 after adjustments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDDSER05(-1)</td>
<td>-5.502811</td>
<td>0.772121</td>
<td>-7.126878</td>
<td>0.0000</td>
</tr>
<tr>
<td>D(DDDSER05(-1))</td>
<td>3.011565</td>
<td>0.662186</td>
<td>4.547914</td>
<td>0.0001</td>
</tr>
<tr>
<td>D(DDDSER05(-2))</td>
<td>1.499508</td>
<td>0.421820</td>
<td>3.554849</td>
<td>0.0015</td>
</tr>
<tr>
<td>D(DDDSER05(-3))</td>
<td>0.557726</td>
<td>0.164849</td>
<td>3.388257</td>
<td>0.0023</td>
</tr>
<tr>
<td>C</td>
<td>-0.917269</td>
<td>0.397029</td>
<td>-1.061941</td>
<td>0.2981</td>
</tr>
<tr>
<td>@TREND’(1970)*</td>
<td>0.572572</td>
<td>0.349893</td>
<td>1.640639</td>
<td>0.1129</td>
</tr>
</tbody>
</table>

R-squared | 0.958335 | Mean dependent var | -3.463216 |
Adjusted R-squared | 0.950322 | S.D. dependent var | 80.59732 |
S.E. of regression | 17.96392 | Akaike info criterion | 8.781968 |
Sum squared resid  | 8390.281 | Schwarz criterion   | 9.056794 |
Log likelihood    | -134.6116  | Hannan-Quinn criter. | 0.973066 |
F-statistic       | 119.6047   | Durbin-Watson stat  | 2.091624 |
Prob(F-statistic) | 0.000000   |                      |           |
The above table shows the cointegration between L1 and T1. There is cointegration in the long run between L1 and T1.

IDENTIFICATION OF THE STOCHASTIC PROCESS

The Box-Jenkins methodology is used to diagnose the stochastic process generating the time series. To find the appropriate stochastic process and the optimal lag lengths we take the help of Correlograms given in figures 1 to 4.

If the autocorrelations taper off slowly or do not die out, non-stationarity is indicated and differencing is suggested until stationarity is obtained. Then an ARMA model is identified for the differenced series.

For an MA(p) process the autocorrelations $a_k=0$ for $k>p$. and the partial autocorrelations taper off. To determine a cut off point of the ACF the sample autocorrelations are used.

For an AR(q) the partial autocorrelations $b_k=0$ for $k>q$ and the autocorrelations taper off. If the spikes of the PACF are significant through $q$ then this determines the degree of AR process.

If neither the autocorrelations nor the partial autocorrelations have a cut off point an ARMA model may be adequate. The AR and the MA degree have to be inferred from the particular patterns of autocorrelations and partial autocorrelations. We find that neither the ACF nor the PACF have a cut off point for the series DL1,DF1, DC1 and DT1. Hence we conclude that the series follow ARMA process. DL1 follows ARMA(2,1), DF1 follows ARMA(2,1) while DC1 follows ARMA(2,1) and DT1 follows ARMA(2,1). These results are rough estimates. They are not supported by any regression analysis.

Cointegration

Cointegration between the time series are studied for estimating a stable long-run equilibrium relationship between the variables concerned. This concept is very useful in empirical analysis because it allows the research to describe the nature of an equilibrium or stationarity relationship between two time series each of which is individually non-stationary.

In our study, we have found that the time series L1,C1,T1 are non stationary at the level, is stationary at the fourth difference, which implies that it is integrated of order three, I(3) and F1 is stationary at the first differencing so it is integrated of order one i.e I(1). In our study, we have taken the series L1, the institutional credit to agricultural sector of India over the time period 1970-2008 as an independent variable, whereas the commercial crop production, C1, food grains production, F1 & the total agricultural production T1, over that time period are taken as the
dependent variables respectively. We want to check, whether there is any longrun stable relationship between the independent variable L1 & each of the dependent variables, C1, F1, T1.

For the study of co-integration between the variables concerned, the following procedures have been adopted-

- The co-integrating equation has been estimated with the OLS method. The non-stationarity of the series under study can be removed by differencing the series, generally the series are integrated of order d, where d>0. Hence the regressand & the regressors of the co-integrating equation are I(d), d>0.
- The residuals of the estimated equation have been obtained. The residuals are the linear combination of the variables which are I(d), included in the equation.
- The residuals are subject to ADF test to examine if random walk exists or if the residuals are white noise, meaning if the residual are I(0).

If the residuals are I(0), then we conclude that the variables are cointegrated otherwise not.

6. CONCLUSION

From the above econometric analysis, we’ve found that there is no co integration of institutional credit with production of foodgrains but cointegration exists with production of commercial crops and total agricultural production. This result has some major economic perspective:

**Firstly,** credit, if considered as an input to agriculture, is the source of monetary capital to the farmers. More specifically, Indian farmers, most of them belonging to low or middle income groups, initially buy capital, raw materials etc with credit because of lack of personal funds. If more credit is issued to the farmers, they can use improved techniques to obtain better produces. So, it can be inferred that Indian agriculture can improve a lot if sufficient amount of credit is issued to agricultural sector and if the issued fund is used efficiently. So higher the credit higher is the agricultural production.

**Secondly,** to be more specific, we’ve considered institutional credit, which means loans issued by commercial banks, regional rural banks and co-operative banks. The obtained non co integration may also result from the dominance of non-institutional credit sources like moneylenders in India. Most of the banks still avoid agricultural sector while issuing credit because of high default rate of farmers. So, the farmers have to depend on non institutional credit sources. Besides, complex credit policies have also refrained the farmers from taking a step towards institutional sources. **Thirdly,** dependence on monsoon and family farming system has created a sort of vicious cycle in Indian agriculture. Family farming process is coupled with disguised unemployment and low productivity. Besides, dependence on monsoon results in fluctuation in agricultural production, which means farmers’ default rate increases in
times of bad monsoon, so banks do not issue credit to them. In turn, lack of fund compels the farmers to stick to old production techniques, behavior of monsoon and non-institutional sources of credit charging high interest rates. So, productivity doesn’t rise significantly.

**REMARKS**

Macro-Economic variables, which are used in this study viz. food grains production (F1), commercial crop production (C1), total agricultural production (P1) and institutional credit (L1) to the agricultural sector, are of time series by nature. These series are not deterministic variables. On the contrary these are considered to be generated by some underlying stochastic processes. Except the series L1 all others show non-stationary at the level. We have taken help of the Box-Jenkins approach to model the time series under study. It has been found that institutional credit (L1) follows an ARMA (2,1) process while foodgrains production (F1), commercial crop production(C1) and total agricultural production(P1 ) follow ARMA processes respectively. The lag lengths for the ARMA processes couldn’t be computed though a rough idea could be obtained by studying the Correlogram. Through Correlogram analysis it has been found that food grains production(F1) follows an ARMA(2,1) process while , commercial crop production (C1) follows an ARMA (2,1) process and total agricultural production (T1) has been generated by the ARMA (2,1) stochastic process.

The time series for food grains production (F1), commercial crop production (C1), total agricultural production (T1) and institutional credit (L1) to the agricultural sector have been subject to tests for stationary. In this study the Augmented Dickey-Fuller method has been adopted for the test of the presence of unit roots for the time series concerned. Unit root tests are undertaken to examine whether the time series exhibit random walk process, i.e. non-stationary. Results such as L1,F1, C1 and P1 have been found to be non stationary Thus L1,F1, C1 and T1 are differenced to make them stationary. L1,C1 and P1 are integrated of order 3 while F1 is I (1).

Non-stationary of the series F1, C1, T1 and L1 at level have further been verified through the estimation the Autocorrelation functions (ACF) and Partial Autocorrelation Function (PACF). The ACF and PACF plots showing the estimated coefficients for different lags along with the upper and lower critical values for the confidence limit have been derived. The plot of ACF and PACF against the lag lengths is known as Correlogram. Through Correlogram analysis it has been verified that L1,F1, C1 and P1 become stationary at fourth difference. In our analysis no co-integration has been found between the time series. There exist no co integration between institutional credit and agricultural food grain production but there exists co integration between institutional credit and total agricultural production.
Institutional credit plays an important role in enhancing the agricultural productivity in developing countries like India. The study discusses about the need for institutional credit followed by a brief about the birth of Institutional credit in India. The paper also discusses the Bank Reforms of India and the impact on the farmers. The study talks about the aversion of the private banks from providing credits to agriculture and the consequences due to it.

It then speaks about the benefits of providing institutional credit for the agricultural sector, but discusses the issues faced by the financial institutions for providing healthy credit. During the study it was found that the Institutional credit has been increasing over the years and there is a direct relation of credit with food grains production. An IMF study over 50 countries ranging from 1980 to 2003 has found evidence of a direct relation between increase in institutional credit and agricultural productivity. The future lies in an alliance between financial institutions and Self-Help groups to provide accessible credit to farmers and provide a win-win situation for the commercial banks.

In India, 70% of the total workforce is employed in agricultural and related sectors. The contribution of agricultural sector towards GDP is 19.9%. Growth of industry and service sectors are also directly linked to agricultural growth. Therefore, in order to achieve a GDP growth of more than 8%, agricultural sector should exhibit a commensurate growth. However, growth of agriculture over the last 10 years has been less than 1.5% per annum. Such a low rate of growth can be attributed to different factors such as low/high monsoon, unavailability of farming equipments and fertilizers, irrigation problems, unavailability of funds etc. institutional credit is required to boost the agricultural sector.

REFERENCES


Das Abhiman, Senapati Manjusha, John Joice, “Impact of Agricultural Credit on Agricultural Production- An Empirical Analysis in India”.


