Forecasting Net Foreign Direct Investment Inflows in India: Box-Jenkins ARIMA Model

Abhijit Biswas

1Dept. of Engg. Science and Humanities, Academy of Technology Adisaptagram, Hooghly, WB, India
2Dept. of Business Administration, University of Kalyani, Kalyani, Nadia, WB, India

Abstract
The study is an attempt to build a time series model to forecast FDI inflows in India over the coming period. Annual time series data for the FDI in India was utilized over the period of 1992-2014. We have collected data mostly from (i) Reserve bank of India and various issues of (ii) Center for Monitoring Indian Economy (CMIE) website. This study employs Regression Analysis, Testing of Parameters, Box Jenkins methodology to build ARIMA (Autoregressive Integrated Moving Average). Accuracy and the selected models were tested by performing different diagnostics tests to ensure the accuracy of the results obtained. There have been enormous forecasting models ranging from simple models to sophisticated ones. We preferred Box – Jenkins model in our project not only due to its simplicity but also for its appropriateness with respect to our sample dataset. The empirical results obtained of ARIMA model have shown that FDI is following an increasing trend over the forecasted period (2015-2034). We used the computer program R GUI, Gretl and Eviews-7 for data analysis and forecasting.

Keywords
FDI, Augmented Dickey Fuller Test, Univariate Analysis, Forecasting, Box-Jenkins methodology, ARIMA, Time Series.

I. Introduction
Foreign Direct Investment (FDI) is defined as the flow of capital from a foreign country to a host country to control assets, establish production or service facilities and to conduct business activities (Park, 2003). Usually 10% stake in equity share and long term continuity in the business in the host country are two important determinants of FDI. Inward FDI has viewed as source of new technology and employment opportunities.

Foreign direct investment plays a vital role in the economic development of the country. It transfers financial resources, technology and innovative and improved management techniques along with raising productivity. An Indian company may receive Foreign Direct Investment either through automatic route or a government route. The paper tries to study the need of FDI in India, to exhibit the sector-wise & year-wise analysis of FDI’s in India, to rank the sectors based upon highest FDI inflows. The results show that Mauritius is the country that has invested highly in India followed by Singapore, Japan, and USA and so on. It also shows that there has been a tremendous increase in FDI inflow in India during the year 2000 to 2011.

FDI have helped India to attain a financial stability and economic growth with the help of investments in different sectors. FDI has boosted the economic life of India and on the other hand there are critics who have blamed the government for ousting the domestic inflows. After liberalization of Trade policies in India, there has been a positive GDP growth rate in Indian economy. Foreign direct investments helps in developing the economy by generating employment to the unemployed, Generating revenues in the form of tax and incomes, Financial stability to the government, development of infrastructure, backward and forward linkages to the domestic firms for the requirements of raw materials, tools, business infrastructure, and act as support for financial system. Forward and back ward linkages are developed to support the foreign firm with supply of raw and other requirements. It helps in generation of employment and also helps poverty eradication. There are many businesses or individuals who would earn their lively hood through the foreign investments. There are legal and financial consultants who also guide in the early stage of establishment of firm.

Foreign investments mean both foreign portfolio investments and foreign direct investments (FDI). FDI brings better technology and management, marketing networks and offers competition, the latter helping Indian companies improve, quite apart from being good for consumers. Alongside opening up of the FDI regime, steps were taken to allow foreign portfolio investments into the Indian stock market through the mechanism of foreign institutional investors. The objective was not only to facilitate non-debt creating foreign capital inflows but also to develop the stock market in India, lower the cost of capital for Indian enterprises and indirectly improve corporate governance structures. On their part, large Indian companies have been allowed to raise capital directly from international capital markets through commercial borrowings and depository receipts having underlying Indian equity. Thus the country adopted a two-pronged strategy: one to attract FDI which is associated with multiple attendant benefits of technology, access to export markets, skills, management techniques, etc. and two to encourage portfolio capital flows which ease the financing constraints of Indian enterprises. Foreign technology induction can be encouraged through FDI and through foreign technology collaboration agreements. The sectors which have resources but do not have the required technology acquire foreign technology collaboration through RBI or Government approvals. The total number of approvals recorded for the period of 2000 to 2010 by the RBI, SIA and FIPB is 8080. The RBI has approved 4580 proposal whereas SIA and FIPB have approved 3500. Technical collaborations have put a positive effect on the domestic firms. It helped in establishing technology transfers. An Indian company may receive Foreign Direct Investment under the two routes as given under:

A. Automatic Route
FDI in sectors /activities to the extent permitted under the automatic route does not require any prior approval either of the Government or the Reserve Bank of India.

B. Government Route
FDI in activities not covered under the automatic route requires prior approval of the Government which are considered by the Foreign Investment Promotion Board (FIPB), Department of Economic Affairs, Ministry of Finance.

II. Literature Review
Factors that influence FDI inflow are (i) intellectual property rights protection (Wu, 2000) and (Javorcik, 2004), (ii) economic
stability and the political climate (Reynolds et al., 2004), (iii) labor market (Giulietti et al., 2004) and (Janicki and Wunnava, 2004), (iv) foreign exchange rate (Klin and (Rosenmagen, 1994)), (v) wages and income convergence (Choi, 2004), (vi) financial and tax policy, (vii) GDP in host country (Shapiro and Globerman, 2003), (viii) bureaucratic corruption and environmental policy (Fredriksson et al., 2003) and so on. Despite the obvious importance of FDI and MNCs in the world economy, researches on the (i) factors that determine FDI patterns and the impact of MNCs on parent and host countries as well as (ii) forecasting of FDI over a span of time that has potential to be useful for policy makers are still in its early stages.

Balasundaram Maniam and Amitava Chatterjee (1998) studied on the determinants of US foreign investment in India; tracing the growth of US FDI in India and the changing attitude of the Indian Government towards it as a part of the liberalization program. Nagesh Kumar (2001) concluded that the magnitudes of inflows have recorded impressive growth, as they are still at a small level compared to the country’s potential. Balasubramaniam VN and Vidya Mahambre (2003) concluded that FDI is a very good means for the transfer of technology and knowhow to the developing countries. Birendra Kumar and Surya Dev (2003) with the data available in the Indian context showed that the increasing trend in the absolute wage of the worker does not deter the increasing flow of FDI. Laura Alfaro (2003) finds that FDI flows into the different sectors of the economy (namely primary, manufacturing, and services) exert different effects on economic growth. FDI inflows into the primary sector tend to have a negative effect on growth, whereas FDI inflows in the manufacturing sector a positive one. Evidence from the foreign investments in the service sector is ambiguous. Sebastian Morris (2004) has discussed the determinants of FDI over the regions of a large economy like India. He argues that, for all investments it is the regions of metropolitan cities that attract the bulk of FDI. Peng Hu (2006) analyses various determinants that influence FDI inflows in India which include economic growth, domestic demand, currency stability, government policy and labour force availability against other countries that are attracting FDI inflows. Analyzing the new findings, it is observed that India has some competitive advantages in attracting FDI inflows, like a large pool of high quality labour force which is an absolute advantage of India against other developing countries like China and Mexico. Chandana Chakraborty and Peter Nuppenkamp (2008) said that booming foreign direct investment in post-reform India is widely believed to promote economic growth. Chew Ging Lee (2009) has pointed out that GDP per capita has a positive effect on FDI inflows in the long run. Krishna Chaitanya Vadlamannatia, Artur Tamazianb and Lokanandha Reddy Iralac (2009) analyses about the determinants of FDI in Asian economies. The determinants are analyzed under four heads, viz. economic and policy factors, socioeconomic factors, institutional factors and political factors. The findings in the baseline models show that poor socioeconomic conditions and labour-related issues are the major determinants. Shiralashetti A.S and S.S. Huger (2009) have made a comparison of FDI inflows during pre and post liberalization period, country-wise, sector-wise and region-wise. Subash Sasidharan and Vinish Kathuria (2011) examine the relationship between FDI and R&D of the domestic firms in the post-liberalization. There are quite a few but noteworthy empirical attempts made by the researchers to examine the growth of FDI inflows using ARIMA models. Here are some of these studies.

Abdel-Rahman (2002) investigates the Determinants of the flow of FDI to the economy of the Kingdom of Saudi Arabia (KSA). The paper discusses FDI with respect to overall trends, sources, and their regional, sectoral and sub-sectoral distributions. It also focuses on the determinants of FDI: the roles of market size, openness and international trade, wage rates, and country risk in attracting FDI to the (KSA). Empirical methods used to gauge the issues include causality tests and conventional regression models where results generally show that activity GDP levels affect FDI in a positive and significant way. Exports had a significant negative impact on the KSA’s FDI, while the socio-political risk variables were mostly significant, and negative in their impacts on FDI inflows.

Karmar and Badkardzhieva (2002) illustrate the reforms needed to attract more FDI investment in Egypt. The paper aims to draw some lessons for Egypt from the experience of Poland, Hungary and the Czech Republic during the 1990s. The paper illustrates that strengthening a country’s attractiveness toward (FDI) has become a new imperative of economic policy. The achievements of the central and eastern European countries in this field appear to be very instructive. The study highlights the importance of multi-regional cooperation as the main determinant of the Egyptian FDI attractiveness.

Shoter and Abdulrazzag (2003) examine the impact of FDI on economic growth in India. The paper illustrates whether or not FDI inflows enhance economic growth in India. It utilizes the augmented production function that includes FDI inflows as the independent variable along with other variables that are expected to have an impact on the growth process. The results show that there is a long-run relationship between economic growth and FDI among other variables.

Alasrag (2005) analyzes development policies of FDI in Arab countries and aims to review and stimulate the mechanisms of FDI flows in the Arab States during 1992-2003. He found out that although many of the reforms that have been taken to attract more FDI in Arab countries, the flow of this investment is still weak compared with other developing countries, such as Mexico, Brazil, Hong Kong and Singapore. The FDI flows in Arab countries were smaller than the flows of FDI in China and the United States (53.5 billion dollars) in the year 2003. On the other hand the investments between the Arab countries in the same period are very low $ 20.7 billion or 44 per cent of the total flows of FDI in Arab countries.


Findings of the study show that ARIMA (0, 1, 1) is the optimal model for forecasting FDI in India and there is an expected increase of FDI volumes over the period (2004-2025).

Judi (2007) uses Autoregressive Integrated Moving Average (ARIMA) models to forecast the non-oil Gross Domestic Product (GDP) in the United Arab Emirates (UAE). The paper analyses the non-oil industry representing the GDP cost prices during the period (1970-2006), which will form a basis to predict future performance of the economy by finding the GDP estimations up to year 2020. That includes the contributions of the different economic sectors other than the oil industry. The main objective of this study is to define the most important sectors in the (UAE) non-oil economy. The outcomes of this study will help in better planning of future strategies, and give an insight of the expected performance of the
economy in the next upcoming fifteen years. Kawaz and Abbadi (2007) identify the risks of FDI on Arab countries. The research aims at knowing the importance and advantages of FDI to the host countries in addition to the risks and determinatives that face FDI in the developing countries, including Arab states, explaining the way each one of these risk is influences on FDI.

In order to obtain empirical results, data was obtained for Arab countries sample and linear regression had been used for the sake of testing the hypothesis of the research. The most important results of the research is that to pay more attention to attracting FDI as it is one of external financial resources also determining the factors affects FDI in order to be increased. Meshaal and Abu Laila (2007) measure and analyze the impact of FDI and imports on economic growth of India, depending on the time series for the period 1976-2003. The study is based on autoregressive (VAR) to achieve this goal, by showing the existence of a causal relationship mutually between FDI, imports and GDP. It found that the same causal relationship between FDI and imports, the existence of indirect effect of foreign investment on human capital, and there is human capital indirect effect on foreign investment through domestic capital and imports.

Sabri (2008) uses the Bayesian models to analyze the impact of FDI on macro-economy in the Republic of Yemen. A forecasting model has been constructed. The importance of the thesis is to design effective and optimal models to develop economic policies which can help attract FDI and also studying the efficiency of predictive models that could assist decision-makers. The findings were that independent variables (budget deficit, the cost of FDI, volume of employment in the investment sector, investment expenses allocated in the state's budget and agricultural production) have significant impact on (GDP) and have a high explanation capacity.

Also FDI has significant impact on exports, imports, agriculture production, extraction industry, manufacturing industry and employment in the investment sector.

Al-Nuemat (2009) explores obstacles and solutions facing FDI in India. He highlights the obstacles facing trans-national corporations (TNC) considering FDI. The paper uses Dunning’s theory which indicates that the third world countries’ ability to attract and make advantage of the potential economic avail from FDI, cultures and infrastructure, differs in accordance with its national, political, economical and legal interests and the government’s policies of the hosting countries together the economical targets. It found that that some of the obstacles encountering FDI in India could probably be attributed to its national infrastructural factors and government policies, as Dunning’s model suggests. The paper recommends improving commercial infrastructure, reinforcing the national competitive capability and the economical policies, raising the economic openness, increasing the government’s investment share in the basic infrastructures, encouraging private sector to join this field and lifting up the level of human resources.

III. Objective

The objective of this study is to forecast the volume of FDI for twenty years (2015 – 2034) beyond the end of sample period (1992-2014). This study employs Box-Jenkins methodology of building ARIMA (Autoregressive Integrated Moving Average) model to achieve the aim of the study. The study is based on Augmented Dickey Fuller test for stationary test, Regression Analysis and the selected models were tested by carry into effect different diagnostic tools to ensure the accuracy of the results obtained. It was found that the time series for the variable (FDI) was not stationary in its level during the time, and it suffers from a unit root, we have been working to make it stationary after identifying first order difference which was used in (ARIMA) models in this study. We used the computer program (R GUI, Gretl and EViews 7.2) for data analysis and forecasting. There have been enormous forecasting models ranging from simple models to sophisticated ones. We have preferred Box-Jenkins model not only due to its simplicity but also for its appropriateness with respect to our sample dataset. The empirical results obtained of ARIMA model have shown that FDI is following an increasing trend over the forecasted period (2015-2034).

IV. Methodology

The study is an attempt to build a time series model to forecast FDI inflows in India over the coming period. Annual time series data for the FDI in India was utilized over the period of 1992-2014. This study employs Regression Analysis, Testing of Parameters, Box Jenkins methodology to build ARIMA (Autoregressive Integrated Moving Average). Accuracy and the selected models were tested by performing different diagnostics tests to ensure the accuracy of the results obtained. There have been enormous forecasting models ranging from simple models to sophisticated ones. We preferred Box – Jenkins model in our project not only due to its simplicity but also for its appropriateness with respect to our sample dataset. We used the computer program R GUI, Gretl and Eviews-7 for data analysis and forecasting.

A. Analysis of Data and Discussion

Table 1:

<table>
<thead>
<tr>
<th>Year</th>
<th>FDI(in Rupees, Crores)</th>
<th>Time</th>
<th>Year</th>
<th>FDI(in Rupees, Crores)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>409</td>
<td>1</td>
<td>2004</td>
<td>10064</td>
<td>13</td>
</tr>
<tr>
<td>1993</td>
<td>1094</td>
<td>2</td>
<td>2005</td>
<td>14653</td>
<td>14</td>
</tr>
<tr>
<td>1994</td>
<td>2018</td>
<td>3</td>
<td>2006</td>
<td>24584</td>
<td>15</td>
</tr>
<tr>
<td>1995</td>
<td>4312</td>
<td>4</td>
<td>2007</td>
<td>56390</td>
<td>16</td>
</tr>
<tr>
<td>1996</td>
<td>6916</td>
<td>5</td>
<td>2008</td>
<td>98642</td>
<td>17</td>
</tr>
<tr>
<td>1997</td>
<td>9654</td>
<td>6</td>
<td>2009</td>
<td>142829</td>
<td>18</td>
</tr>
<tr>
<td>1998</td>
<td>13548</td>
<td>7</td>
<td>2010</td>
<td>123120</td>
<td>19</td>
</tr>
<tr>
<td>1999</td>
<td>12343</td>
<td>8</td>
<td>2011</td>
<td>97320</td>
<td>20</td>
</tr>
<tr>
<td>2000</td>
<td>10311</td>
<td>9</td>
<td>2012</td>
<td>165146</td>
<td>21</td>
</tr>
<tr>
<td>2001</td>
<td>10733</td>
<td>10</td>
<td>2013</td>
<td>121907</td>
<td>22</td>
</tr>
<tr>
<td>2002</td>
<td>18654</td>
<td>11</td>
<td>2014</td>
<td>147518</td>
<td>23</td>
</tr>
<tr>
<td>2003</td>
<td>12871</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the Table 1 given above we have to test the stationary stage of time series i.e. whether it has a unit root or not or in other words we have to perform unit root test for stationary of time series.

1. Unit Root Test

At first let us define non - stationary. Considering the simplest stochastic trend model as :-

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

or,

\[ \Delta y_t = u_t \]

Generalising the concept to consider the case where the series contains more than one “unit root”. We need to apply the first difference operator, \( \Delta \), more than once to induce stationary.
If a non-stationary series, $y_t$, is differenced ‘d’ times before it becomes stationary then it is said to be integrated of order ‘d’. We write $y_t \sim I(d)$. So if $y_t \sim I(0)$ then $\Delta y_t \sim I(0)$. An I(0) series is a stationary series. An I(1) series contains one unit root.

The tests are based on the $t$-ratio on $y_{t+1}$ against a trend.

- Case (i): $y_t = \phi y_{t+1} + u_t$
- Case (ii): $y_t = \phi y_{t+1} + \mu + \lambda t + u_t$
- Case (iii): $y_t = \phi y_{t+1} + \mu + \lambda + u_t$

In each case, the tests are based on the $t$-ratio on $y_{t+1}$ against a trend. This is a test for a random walk against a stationary AR(1) with $\phi < 1$.

The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (Dickey and Fuller 1979, Fuller 1976).

**Dickey Fuller tests** are also known as τ tests: $\tau$, $\tau_1$, $\tau_2$, $\tau^*$. Dickey Fuller tests are also known as τ tests: $\tau$, $\tau_1$, $\tau_2$. $\tau^*$ models in each case.

### Null Hypothesis:
- $H_0$: series contains a unit root
- $H_1$: series is stationary.

And we usually use the regression:

$$\Delta y_t = \psi y_{t+1} + u_t$$

so that test of $\phi = 1$ is equivalent to test of $\psi = 0$ (since $\phi - 1 = \psi$).

**Dickey Fuller tests** are also known as τ tests: $\tau$, $\tau_1$, $\tau_2$. $\tau^*$勇 for constant $\mu$ and trend $\lambda$.

By computing the Dickey Fuller Test we can write:

$$\Delta y_t = u_t$$

where $\Delta y_t = y_{t+1} - y_t$, and the alternatives may be expressed as:

- $y_t = \psi y_{t+1} + \mu + \lambda + u_t$ (where $\mu = \lambda = 0$ in case i).
- $y_t = \psi y_{t+1} + \mu + u_t$ (where $\lambda = 0$ in case ii.) and $\psi = \phi - 1$.

**[N.B.](The null ($H_0$) and alternative ($H_1$) models in each case are)

- i.) $H_0$: $y_t = y_{t+1} + u_t$
- $H_1$: $y_t = \phi y_{t+1} + u_t$  \(\phi < 1\)

This is a test for a random walk against a stationary autoregressive process of order (AR(1)).

- ii.) $H_0$: $y_t = y_{t+1} + u_t$  \(\phi < 1\)
- $H_1$: $y_t = \phi y_{t+1} + \mu + u_t$  \(\phi < 1\)

This is a test for a random walk against a stationary AR(1)  \(\phi < 1\) with drift.

- iii.) $H_0$: $y_t = y_{t+1} + u_t$
- $H_1$: $y_t = \phi y_{t+1} + \mu + \lambda + u_t$. \(\phi < 1\)

This is a test for a random walk against a stationary AR(1) with drift.

### Decision Rule:
- If $t^* > ADF$, PP critical value, the decision: reject null hypothesis, unit root does not exist.
- If $t^* < ADF$, PP value (in absolute terms) < $t^*$ critical value (in absolute terms), the decision: not reject null hypothesis, unit root exists.

### Phillips - Perron Test:
Phillips and Perron have developed a more complicated theory of unit root non-stationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow auto correlated residuals. The tests usually gives the same conclusions as the ADF tests and the calculation of test statistic is complex.

### Phillips - Perron Test:
- If $t^* > ADF$, PP critical value, the decision: not reject null hypothesis, unit root exists, the value series is non-stationary.
- If $t^* < ADF$, PP critical value, the decision: reject null hypothesis, unit root does not exist, the value series is stationary.

### Critical values for DF and ADF tests (Fuller, 1976, P373)
The null hypothesis of a unit root is rejected in favor of the stationary alternative in each case if the test statistic is more negative than the critical value.

### The test above are only valid if $u_t$ is white noise. In particular, $u_t$ will be autocorrelated if there was autocorrelation in the dependent variable of the regression ($\Delta y_t$) which we have not modelled. The solution is to “augment” the test using $p$ lags of the dependent variable. The alternative model in case (i) is now written as:

$$\Delta y_t = \psi y_{t+1} + \sum_{i=1}^{p} \eta_i \Delta y_{t-i} + u_t$$

The same critical values from the DF tables are used as before. A problem arises now in determining the optimal number of lags of the dependent variable.

There are 2 possible ways:
- By using the frequency of the data to decide.
- By using information criteria.

### Phillips - Perron Test:
Phillips and Perron have developed a more complicated theory of unit root non-stationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow auto correlated residuals. The tests usually gives the same conclusions as the ADF tests and the calculation of test statistic is complex.

### Phillips - Perron Test:
- If $t^* > ADF$, PP critical value, the decision: reject null hypothesis, unit root exists, the value series is non-stationary.
- If $t^* < ADF$, PP critical value, the decision: not reject null hypothesis, unit root does not exist, the value series is stationary.

### Table 2: ADF Test for FDI in its level.

<table>
<thead>
<tr>
<th>Lag Length: 0 (Automatic - based on AIC, maxlag=3)</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-2.084286</td>
<td>0.5257</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-4.440739</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.632896</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.254671</td>
<td></td>
</tr>
</tbody>
</table>


**Augmented Dickey-Fuller Test Equation**
Dependent Variable: D(FDI)
Method: Least Squares
Date: 08/23/15 Time: 12:02
Sample (adjusted): 1993 2014
Included observations: 22 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>FDI(-1)</td>
<td>-0.351068</td>
<td>0.168435</td>
<td>-2.084286</td>
<td>0.0509</td>
</tr>
<tr>
<td>C</td>
<td>-3.181981</td>
<td>0.164647</td>
<td>-1.84730</td>
<td>0.0627</td>
</tr>
<tr>
<td>@TREND(1992)</td>
<td>3111.774</td>
<td>1392.897</td>
<td>2.234030</td>
<td>0.0377</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.212844</td>
<td>6686.773</td>
<td>6686.773</td>
<td>6686.773</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.129985</td>
<td>24113.99</td>
<td>24113.99</td>
<td>24113.99</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>2249.23</td>
<td>2249.23</td>
<td>2249.23</td>
<td>2249.23</td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>9.61E+09</td>
<td>2249.23</td>
<td>2249.23</td>
<td>2249.23</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-250.0644</td>
<td>2249.23</td>
<td>2249.23</td>
<td>2249.23</td>
</tr>
<tr>
<td>F-statistic</td>
<td>2.568758</td>
<td>2.249922</td>
<td>2.249922</td>
<td>2.249922</td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.102939</td>
<td>0.102939</td>
<td>0.102939</td>
<td>0.102939</td>
</tr>
</tbody>
</table>

### Table 2: Critical values for the DF Test

<table>
<thead>
<tr>
<th>Significance Level</th>
<th>10 %</th>
<th>5 %</th>
<th>1 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C.V$ for constant but no trend</td>
<td>-2.57</td>
<td>-2.86</td>
<td>-3.43</td>
</tr>
<tr>
<td>$C.V$ for constant and trend</td>
<td>-3.12</td>
<td>-3.41</td>
<td>-3.96</td>
</tr>
</tbody>
</table>

Critical values for DF and ADF tests (Fuller, 1976, P373)
The computed ADF test statistics is greater than critical values at different levels so we cannot conclude to reject null hypothesis. FDI series has unit root problem and the series is non stationary.

Table 3: Phillips- Perron test for FDI in its level

Null Hypothesis: FDI has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 0 (Newey-West automatic) using Bartlett kernel

<table>
<thead>
<tr>
<th>Adj. t-Stat</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phillips-Perron test statistic</td>
<td>-2.084286 0.5257</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-4.440739</td>
</tr>
<tr>
<td>5% level</td>
<td>-3.632896</td>
</tr>
<tr>
<td>10% level</td>
<td>-3.254671</td>
</tr>
</tbody>
</table>


Residual variance (no correction) 4.37E+08
HAC corrected variance (Bartlett kernel) 4.37E+08

ADF test Equation
Dependent Variable: D(FDI)
Method: Least Squares
Date: 08/23/15   Time: 12:17
Sample (adjusted): 1993 2014
Included observations: 22 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>FDI(-1)</td>
<td>-0.351068</td>
<td>0.168435</td>
<td>-2.084286</td>
<td>0.0509</td>
</tr>
<tr>
<td>C</td>
<td>-13818.91</td>
<td>11664.19</td>
<td>-1.184730</td>
<td>0.2507</td>
</tr>
<tr>
<td>@TREND(1992)</td>
<td>3111.774</td>
<td>1392.897</td>
<td>2.234030</td>
<td>0.0377</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.212844</td>
<td>Mean dependent var</td>
<td>6686.773</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.129985</td>
<td>S.D. dependent var</td>
<td>2411.99</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>22492.23</td>
<td>Akaike info criterion</td>
<td>23.00585</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>9.61E+09</td>
<td>Schwarz criterion</td>
<td>23.15463</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-250.0644</td>
<td>Hannan-Quinn criter.</td>
<td>23.04090</td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>2.568758</td>
<td>Durbin-Watson stat</td>
<td>2.249922</td>
<td></td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.102939</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the above table the absolute value of ADF test statistic is greater (in absolute terms) than the critical values at different significant levels respectively, thus we can conclude that it is the first differences of the value series that are stationary. Thus the first difference time series can be used for forecasting. The Durbin Watson statistics is significant at 2.026413.

Table 5: Phillips - Perron Unit Root Test for FDI with First Difference

Null Hypothesis: D(FDI) has a unit root
Exogenous: Constant
Bandwidth: 0 (Newey-West automatic) using Bartlett kernel

<table>
<thead>
<tr>
<th>Adj. t-Stat</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phillips-Perron test statistic</td>
<td>-5.791684 0.0001</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-3.788030</td>
</tr>
<tr>
<td>5% level</td>
<td>-3.012363</td>
</tr>
<tr>
<td>10% level</td>
<td>-2.646119</td>
</tr>
</tbody>
</table>


Residual variance (no correction) 5.32E+08
HAC corrected variance (Bartlett kernel) 5.32E+08

ADF test Equation
Dependent Variable: D(FDI,2)
Method: Least Squares
Date: 08/23/15   Time: 12:51
Sample (adjusted): 1994 2014
Included observations: 21 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(FDI(-1))</td>
<td>-1.290890</td>
<td>0.222887</td>
<td>-5.791684</td>
<td>0.0000</td>
</tr>
<tr>
<td>C</td>
<td>8655.552</td>
<td>5446.347</td>
<td>1.589240</td>
<td>0.1285</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.638396</td>
<td>Mean dependent var</td>
<td>1186.952</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.619364</td>
<td>S.D. dependent var</td>
<td>3930.56</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>24248.62</td>
<td>Akaike info criterion</td>
<td>23.12050</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>1.12E+10</td>
<td>Schwarz criterion</td>
<td>23.21998</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-240.7652</td>
<td>Hannan-Quinn criter.</td>
<td>23.14209</td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>33.54360</td>
<td>Durbin-Watson stat</td>
<td>2.026413</td>
<td></td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.000014</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the above table the absolute value of ADF test statistic is greater (in absolute terms) than the critical values at different significant levels respectively, thus we can conclude that it is the first differences of the value series that are stationary. Thus the first difference time series can be used for forecasting. The Durbin Watson statistics is significant at 2.026413.

2. Box - Jenkins Model (ARIMA)

The Box-Jenkins approach to modelling Auto Regressive, Integrated, and Moving Average (ARIMA) processes is a
mathematical model used for forecasting. Box-Jenkins modelling involves identifying an appropriate ARIMA process, fitting it to the data, and then using the fitted model for forecasting. One of the attractive features of the Box-Jenkins approach to forecasting is that ARIMA processes are a very rich class of possible models and it is usually possible to find a process which provides an adequate description to the data. The original modelling procedure involved an iterative three-stage process of model selection, parameter estimation and model checking ((Box and Jenkins, 1970; Vandaele, 1983).

Recent explanations of the process often add a preliminary stage of data reparation and a final stage of model application or forecasting. Each ARIMA process has three parts: the autoregressive (or AR) part; the integrated (or I) part; and the moving average (or MA) part. The models are often written in shorthand as ARIMA (p, d, and q) where p describes the AR part, d describes the integrated part and q describes the MA part.

AR: This part of the model describes how each observation is a function of the previous p observations. For example, if p = 1, then each observation is a function of only one previous observation. That is,

\[ Y_t = c + \phi Y_{t-1} + e_t \]

where \( Y_t \) represents the observed value at time t, \( Y_{t-1} \) represents the previous observed value at time \( t-1 \), \( e_t \) represents some random error and \( c \) and \( \phi \) are both constants. Other observed values of the series can be included in the right-hand side of the equation if \( p > 1 \):

\[ Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t. \]

(I): This part of the model determines whether the observed values are modelled directly, or whether the differences between consecutive observations are modelled instead. If \( d = 0 \), the observations are modelled directly. If \( d = 1 \), the differences between consecutive observations are modelled. If \( d = 2 \), the differences of the differences are modelled. In practice, \( d \) is rarely more than 2.

MA: This part of the model describes how each observation is a function of the previous q errors. For example, if q = 1, then each observation is a function (Makridakis, Wheelwright and Hyndman, 1998).

Although we have made a unit root test, and prove that FDI series was not stationary, and determined the first difference for ARIMA models, but it is better to emphasize it again before we go ahead with forecasting.

(Ref:- www.ccsenet.org/iijbm , Vol. 6. 10; October 2011)

3. Data Preparation

From the line graph in Figure (1), we can see that the time series is likely to have upward trend and seasonal cycles, which implies to non - stationary level. It is clear that variance in the FDI series is not stable where the variation changes with the level, an indication that is not stationary. This means that the short term mean level is not constant but varies over the time series. Similarly the trend in Forecast has a upward trend from figure given below. It has been taken into account the first difference while representing the graph.

![Fig. 1: Plot of FDI in Rs. (Cr’s) of Indian Rupees With Forecasts](image1)

![Fig. 2: Autocorrelation (ACF) and Partial Autocorrelation Function (PACF) of FDI](image2)

Autocorrelation Function (ACF) is used as shown in fig. 2 below. It illustrate that there is a significant spike at ACF at lag 1, and after the first lag, the ACFs are slowly declined. We can conclude again that time series is non - stationary. Again from the fig. 2 below Partial Autocorrelation Function (PACF) of the difference series in the estimation period, we see that it has a significant spike at lag 1. The mean and variance do not remain constant throughout the time periods indicating the non-stationarity of the time series. Since the ACF and PACF have spikes at lag 1, so the differences can be used for ARIMA model.
Graph in fig. 3 shows that after taking differences the time series became stationary, noting that the variance is stable where the variation changes with the level.

Fig. 3: Time series plot of FDI after taking differences in Rupees in Crores(INR)

From fig. 4 below, it is clearly that the ACF first difference series has no significant spikes at any lags. We can conclude that time series is stationary.

Fig. 4: ACF and PACF After Taking Differences

4. Estimating ARIMA Models
Since the time series become stationary after the first difference, it is possible to estimate the following models and choose the most appropriate model for forecasting. The autoregressive (or AR) part; the integrated (or I) part; and the moving average (or MA) part. The models are often written in shorthand as ARIMA (p, d, and q) Where p describes the AR part, d describes the integrated part and q describes the MA part.

5. Parameter Estimation
In order to find the values of the model coefficients which provide the best fit to the data? And testing the assumptions of the model to identify any areas where the model is inadequate.

We suggest using the first difference included in the following models:-
- The autoregressive model AR (1, 1, 0) - the moving average model MA (0, 1, 1)
- The integrated (AR) part; and (MA) part. Written as ARIMA (1, 1, 1)

Autoregressive model AR (1, 1, 0)

Table 6:
AR (I) Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Cochrane-Orcutt, using observations 1994-2014 (T = 21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable: d_Fdi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rho</td>
<td>-0.29089</td>
<td>3995.3</td>
<td>0.1089</td>
<td></td>
</tr>
</tbody>
</table>

Statistics based on the rho-differenced data:
- Mean dependent var: 6972.571
- S.D. dependent var: 24671.28
- Sum squared resid: 1.12e+10
- S.E. of regression: 23634.63
- R-squared: 0.000000
- Adjusted R-squared: 0.000000
- Durbin-Watson: 2.026413

From the Table 6 the coefficient of AR(1) with first difference is not statistically significant at (0.05) level as the value of P is greater than 0.05, thus we ignore this model.

Similarly it gives the same result for MA(1).

In the fig. 5 below ACF and PACF of residuals shows that FDI series has no problem with residuals and there are no spikes which indicate a good sign for using this model for forecasting.

So its likely to use (0, 1, 1) model for forecasting. Autoregressive model AR and moving average model MA with first difference ARIMA(1, 1, 1).
Table 7:

<table>
<thead>
<tr>
<th>ARIMA (1, 1, 1) statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Estimates of Parameters</td>
</tr>
<tr>
<td>Type</td>
</tr>
<tr>
<td>AR 1</td>
</tr>
<tr>
<td>MA 1</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

Differencing: 1 regular difference

Number of observations: Original series 22, after differencing 21

Residuals: SS = 1177003293 (back forecasts excluded)

MS = 653890739 DF = 18

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>3.0</td>
<td>9</td>
<td>0.964</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Results obtained from Table 7 states that AR(1) with the first difference is not statistically significant at (0.05) level as “P” value (0.195) is greater than (0.05), thus we ignore the model.

ARIMA(0, 1, 1) statistics

Model 1: ARIMA, using observations 1994-2014 (T = 21)
Dependent variable: (1-L) d_Fdi

Standard errors based on Hessian

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>673.063</td>
<td>797.104</td>
<td>0.8444</td>
</tr>
<tr>
<td>theta_1</td>
<td>-0.999999</td>
<td>0.127364</td>
<td>-7.8515</td>
</tr>
</tbody>
</table>

Mean dependent var 1186.952 S.D. dependent var 39303.56
Mean of innovations -81.03726 S.D. of innovations 23714.44
Log-likelihood -242.8939 Akaike criterion 491.7877
Schwarz criterion 494.9213 Hannan-Quinn 492.4678

ARIMA(1, 1, 1) statistics

Model 1: ARIMA, using observations 1994-2014 (T = 21)
Dependent variable: (1-L) d_Fdi

Standard errors based on Hessian

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-118.031</td>
<td>3876.21</td>
<td>-0.03045</td>
</tr>
<tr>
<td>phi_1</td>
<td>-0.651568</td>
<td>0.172082</td>
<td>-3.786</td>
</tr>
</tbody>
</table>

Mean dependent var 1186.952 S.D. dependent var 39303.56
Mean of innovations 15.39437 S.D. of innovations 29695.17
Log-likelihood -246.3475 Akaike criterion 491.7877
Schwarz criterion 501.8286 Hannan-Quinn 499.3751

ARIMA(1, 1, 0)

ARIMA, using observations 1994-2014 (T = 21)
Estimated using Kalman filter (exact ML)
Dependent variable: (1-L) d_Fdi

Standard errors based on Hessian

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>std. error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-118.031</td>
<td>3876.21</td>
<td>-0.03045</td>
</tr>
<tr>
<td>phi_1</td>
<td>-0.651568</td>
<td>0.172082</td>
<td>-3.786</td>
</tr>
</tbody>
</table>

Mean dependent var 1186.952 S.D. dependent var 39303.56
Mean of innovations 15.39437 S.D. of innovations 29695.17
Log-likelihood -246.3475 Akaike criterion 491.7877
Schwarz criterion 501.8286 Hannan-Quinn 499.3751

ARIMA(0, 1, 1) statistics

Model 1: ARIMA, using observations 1994-2014 (T = 21)
Dependent variable: (1-L) d_Fdi

Standard errors based on Hessian

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>const</td>
<td>-118.031</td>
<td>3876.21</td>
<td>-0.03045</td>
</tr>
<tr>
<td>phi_1</td>
<td>-0.651568</td>
<td>0.172082</td>
<td>-3.786</td>
</tr>
</tbody>
</table>

Mean dependent var 1186.952 S.D. dependent var 39303.56
Mean of innovations 15.39437 S.D. of innovations 29695.17
Log-likelihood -246.3475 Akaike criterion 491.7877
Schwarz criterion 501.8286 Hannan-Quinn 499.3751

6. Model Checking

Table 8: (Values of the model co-efficients.)

ARIMA(1, 1, 1)
Forecast evaluation statistics

Mean Error 21.691
Mean Squared Error 5.3828e+008
Root Mean Squared Error 23201
Mean Absolute Error 15409
Mean Percentage Error 41.546
Mean Absolute Percentage Error 173.65
Theil’s U 0.56955

ARIMA(0, 1, 1)
Forecast evaluation statistics

Mean Error -81.037
Mean Squared Error 5.9451e+008
Root Mean Squared Error 24383
Mean Absolute Error 15981
Mean Percentage Error 57.492
Mean Absolute Percentage Error 152.1
Theil’s U 0.58459

ARIMA(1, 1, 0)
Forecast evaluation statistics

Mean Error 15.394
Mean Squared Error 8.1811e+008
Root Mean Squared Error 29695
Mean Absolute Error 17696
Mean Percentage Error 140.08
Mean Absolute Percentage Error 141.31
Theil’s U 1.2364

From the above values of model it is good to keep the third model for consideration.

Table 9: Forecasting results for the period 2015-2034
Forecasts from period 23 with difference.

For 95% confidence intervals, z(0.025) = 1.96
The models can be made more appropriate if the data points are about forecasting FDI in India. This study will encourage researchers to conduct further studies as it provides several important insights into issues relating to forecasting. Hopefully, accordingly. Despite this limitation, this study has provided several insights into issues relating to forecasting. Thus could be varied from the time series and any added data could change the results. Box Jenkins methodology for forecasting. While ARIMA models a limitation and might be explored in future research. It adopted 6 - It's very essential for decision –makers to take a look at the results of this research. 7 - This study is subject to a limitation and might be explored in future research. It adopted box Jenkins methodology for forecasting. While ARIMA models limiting the choice of methodology, which is only employing time series data collection for forecasting. Thus could be varied from one study to another one that depends on the number of years used in the time series and any added data could change the results accordingly. Despite this limitation, this study has provided several important insights into issues relating to forecasting. Hopefully, this study will encourage researchers to conduct further studies about forecasting FDI in India. The models can be made more appropriate if the data points are monthly or quarterly- further research.

V. Conclusion
Table (9) represents the conclusion of the study which is the forecasting FDI over the coming twenty years and we found out the following findings:- The total volume of direct investment is expected for the years (2015-2034) is (276130) rupees crores in INR. There is an expected smooth increase of FDI inflows to India over the years (2015-2034).

VI. Recommendations
To provide a 1-suitable investment environment in India through more incentives and facilities to investors away from the bureaucracy and the removal of the obstacles faced by these investors. 2- To work on creating investment opportunities to attract more FDI in the country. – To work on a comprehensive economic plan, creating more job opportunities, reducing poverty. 4 - Conduct a comprehensive review of all legislation governing FDI in general and the Investment Promotion Law in particular. 5 - To pay more efforts by the government on fighting all forms of corruption. 6- It’s very essential for decision-makers to take a look at the results of this research. 7 - This study is subject to a limitation and might be explored in future research. It adopted box Jenkins methodology for forecasting. While ARIMA models limiting the choice of methodology, which is only employing time series data collection for forecasting. Thus could be varied from one study to another one that depends on the number of years used in the time series and any added data could change the results accordingly. Despite this limitation, this study has provided several important insights into issues relating to forecasting. Hopefully, this study will encourage researchers to conduct further studies about forecasting FDI in India. The models can be made more appropriate if the data points are monthly or quarterly- further research.

References


[17] The use of Bayesian models to analyze the impact of FDI on the overall economy of Yemen.


[21] Softwares Used:- GRETL, MINITAb, EVIEWS-7, R


An astute MBA (Systems) & M.Sc. (Economics with Statistics & Econometrics as specialization), B.Sc in Eco-Stat, IRDA Trained, SMDA from VGSOM, IIT Kharagpur, Ph.D student, Data Analytics certified professional with 12 years of experience in Teaching, Educational Planning, Administration, Consultancy and Training & Development across Education sector. A keen planner & implementer with track record of implementing operational policies / norms, systems & controls, motivational schemes & education standards for professionals during the career span. Actively involved in supervising and managing planning and development as well as stellar in imparting knowledge to students on Management, OR, Statistics, Economics and Computer. Started career as Assistant Professor & Teacher In-charge Training with Academy of Technology, Hooghly, West Bengal and gradually rose to a level of Assistant Professor with the same organisation. Visiting Professor in the Department of Business Administration in Kalyani University. Demonstrated strong abilities in assessing organizational development needs, develop solutions for critical enterprise level initiatives and provide mentorship and guidance for excellent solution implementation.