CHAPTER III

METHODOLOGY

The methodology for the present study is presented and discussed under the following heads:

I. Background of the study.
II. Sample Design of the study
III. Data base of the study
IV. Tools of analysis
V. Limitations of the study.

I. Background of the Research Study

The Indian banking industry has come from a long way from being passive business institution to a highly proactive and dynamic entity. Before liberalisation, the Indian banking structure was largely controlled and parameters like branch size and location were given much importance. Presently, the Indian banking industry is going through a period of intense change, where liberalized business environment has affected the banking business by way of increasing competition, rising customer expectations, shrinking spreads and increasing disintermediation. Ongoing changes in the structure of Indian banking are clearly visible. While the share of public sector banks in the total assets of the banking sector has shown a steady decline, the new private sector banks and foreign banks have succeeded in enhancing their position. As compared to the techequipped distribution network of the new private sector banks and foreign banks, public sector banks have found it difficult to upgrade themselves on the technology front. Public sector banks have also started embracing technology to improve customer service and design innovative products to increase sales opportunities and have started undertaking various cost-reduction programmes.

In the face of growing competition, the policy changes and the operational environment in respect of the Indian banking system, there has been an increased focus on profitability, although other social objectives continue to be important. The profit performance has been quite varied among different bank groups and within each group in respect of individual banks as well. Generally, new private sector banks and foreign banks have performed better than public sector banks and old private sector banks. While the level of non-performing assets of public sector banks remains high, a noteworthy development has been their significant reduction in relation to net advances in recent years. In order to remain in the competition, it is required to convert the challenge of change into exciting opportunity. The performance of banks has become a major
concern to planners and policy makers in India, since, the gains of real sector economy depend on how efficiently the financial sector performs the function of financial intermediation.

The present study is an attempt to examine the impact of liberalization on the productivity and profitability of commercial banks in India. The Indian banking sector is of particular interest for a number of reasons. First of all, the pragmatic and gradual philosophy of the Indian deregulation policy offers a great scope for analysing the relationship between the extent of regulatory constraints and the efficiency and productivity of economic agents. Secondly, the diverse ownership of the Indian banking system provides an opportunity for a test of performance differentials between public, private and foreign banks in reaction to a changing regulatory environment. Thirdly, although this study relates to the Indian experience, it has a broader appeal. Indian banking is a considerable component of Asian financial markets and it shares quite similar characteristics with the banking system of other Asian countries. Given the fact that most countries located in the Asian region have either been embarking on a deregulation path or are contemplating to initiate deregulation, an empirical investigation of the deregulation effects on efficiency and the dynamics of productivity change in the Indian case could provide useful policy suggestions to those countries.

Further, most of the earlier studies looked only into the productivity aspect of performance but not on the other aspects viz., profitability, financial management and asset quality, which were focused by the post-liberalization committee’s recommendations. In the light of the above discussions, the objective of the present study was focused on to estimate the efficiency of commercial banks operating in India with five indicators i.e., productivity, operations, profitability, financial management and asset quality. This study contributes to the existing efficiency and productivity literature in several ways. Firstly, we use a time period which encompasses the whole deregulation process. This enables us to perform a comprehensive analysis of the processes of adjustment and adaptation of the Indian banking system to a new regulatory structure. Secondly, the study uses a broad dynamic indicator, total factor productivity (TFP) growth index, as the measure of performance. Ahluwalia and Shanker (1985) pointed out that one needs to look at the concept of TFP to evaluate the performance of the service sector in India, which not merely reflects productive efficiency of labour and capital but also the capacity of the management to combine these and other factors to enhance the output of the firm. In this study, TFP growth was calculated using a DEA-type Malmquist index and further decompose the measured TFP growth into technological change and technical efficiency change, which enables one capture information about the contributory elements of productivity
improvement. Both the quantitative change in productivity and the sources of the change have important policy implications and are relevant in examining the reform effect.

II. Sample Design of the Study

The banking sector in India comprises of domestic banks, which includes public sector banks and privatively owned banks as well as foreign banks. Presently, there are 43 Foreign Banks, 26 Public Sector Banks and 20 Private Sector Banks (RBI, 2013). The number of public sector commercial banks has almost remained the same over last three decades. And in terms of asset share, the public sector banks constitute about 70 percent of the total commercial banking asset. But the point to be noted, the asset share of the public sector banks has gone down from about 90 percent in 1980 to about 68 percent in 2007 and had a gradual increase since 2009 and was 73 percent in 2013. Even though the number of domestic private banks has declined since 1980s, the asset share of these banks has gone up to about 20 percent in 2007 and maintained stability over the years (21 percent in 2013). On the other hand, even though the number of foreign banks has gone up significantly, their asset share has not increased in a significant way. These different types of banks differ from each other in terms of operational efficiency, productivity, profitability and credit efficiency. The study excluded the regional rural banks, since these banks were established for providing credit to local bodies for social development at minimum interest rate and their inclusion may lead to misleading conclusions. The objective of the study was to identify the differences within the banking groups. To make a comparative analysis of various indicators the banks were divided into four groups, i.e., All Scheduled Commercial Banks, Public Sector Banks, Private Sector Banks, and Foreign Banks.

III. Data base of the study

The entire structure of data for the study rests solely on secondary sources of information. The study was carried out during the period from 1980–81 to 2012–13. To examine whether the relationship was stable across the sample period or whether there were any significant changes in the relationship in the latter parts of 1990s the sample period was divided into: pre-period (1981–1992) and post-period (1993–2013). The significance of classification of the time periods was that though the recommendations of Narasimham Committee Report were adopted by the end of 1991, its effect was more discernible only since 1993. Hence the study had taken 1993 as the beginning of the post-reform period assuming a lag of one year for the implementation of the reforms. The research was based on the secondary data and the main stream of data was collected from the “Statistical Tables Relating to Commercial Banks in India”, ‘Annual Accounts of Scheduled Commercial Banks’, Annual Reports, authentic records
Data on GDP was collected from ‘Handbook of Statistics on the Indian Economy’ published by the Reserve Bank of India. All the variables have been adjusted to the Indian GDP deflator at constant prices (2004–05 prices).

IV. Tools of analysis

Since the research was based on time series cross section secondary data and quantitative in nature, the analysis of the data was undertaken by applying the analytical tools like ratio analysis, trend analysis, DEA analysis, productivity indices, and regression and correlation analysis. Before doing the computations, the stationarity of each series was tested using the Augmented Dickey–Fuller test. The conclusions were drawn on the basis of 5 percent level of significance.

Ratio Analysis:

Ratio analysis is quite reliable and provides relationships amongst the various items in the balance sheet of banking system, which can be further used to draw conclusions. The ratios also provide a convenient means of analysis and expression of the various operational aspects of banking business. It is a very useful diagnose tool for assessing the performance of different ownership in time series data. The ratios which constitute the five indicators are as follow:

A. Productivity Efficiency

The production process in banking involves the use of deposits and other assets. Deposit mobilization and advances are two major business products of every banking organization. In fact, productivity indicated by per employee, per office and some financial indicators are popular all over the world and are mostly relied upon. Productivity ratios of total deposits, total loans and total business have been calculated in relation to total manpower employed and number of business centres (branches), as detailed below:

i. Business per Employee = Deposit + Advances / No. of Employees

ii. Advances per Employee = Total Advances / No. of Employees

iii. Deposits per Employee = Total Deposit / No. of Employees

iv. Deposit per Branch = Total Deposit / No. of Branches.

v. Advances per Branch = Total Advances / No. of Branches

vi. Business per Branch = Deposit + Advances / No. of Branches

B. Operations Efficiency

The ratios explain the operational success of the bank. The result will enable one to know how well the bank is running and whether it is able to sustain its stability for the further
earning of the business. The bank becomes insolvent if it’s unable to meet its expenses and achieve proper utilisation of its average working funds. The operational efficiency of the banks was judged on the basis of the following ratios:

i. Interest Income / Average Working Funds  
ii. Non-Interest income / Average Working Funds  
iii. Operating Expenses / Operating Income  
iv. Cost of Deposits = Interest Paid on Deposits / Average Deposits.  
v. Spread / Average Working Funds  
vi. \( \text{CAR} = \frac{\text{Tier I capital} + \text{Tier II capital}}{\text{Risk Weighted Assets}} \) (reliable form of capital)

C. Profitability Efficiency  
The profitability indicates the ability of the firm to generate earnings. The ratio reveals the difference between the revenue and the cost. The expenses are cost incurred to generate revenue but if cost overhead the revenue there arise the question of profit. Hence these ratios explain how well the bank uses its earnings and manage the expenses and attain profitability. These calculations facilitate the management to identify the surpassed expenses and bestow the cost control measures. These criterions of profitability are judged on the basis of the following ratios:

i. Operating Profit Margin  

\[ \text{OPM} = \frac{\text{Interest Income} - \text{Operating Expenses}}{\text{Total interest income}} \]

ii. Net Profit Margin = Net Profit / Total Income  
iii. Operating profit / Average Working Funds  
iv. Return on capital employed = Net Profit after Tax / Average Working Funds  
v. Credit / Deposit Ratio

D. Asset Quality  
The asset quality of banks is mainly replicated in the performance of asset to that of credit lending resources with and without provisions made on the performance of asset. This ratio will help the bank to assess its efficiency of whether it has utilized its resources in a productive manner and whether it is able to withhold or strengthen its capital position irrespective of unexpected losses. The non performance of the asset will raise losses of the bank which indicates poor management, while lower ratio reflects the structural consistency of the bank. The ratios used were:

i. Gross NPAs / Gross Advances
ii. Gross NPAs / Total Assets

iii. \( \frac{\text{Net NPAs}}{\text{Net Advances}} = \frac{\text{Gross NPAs} - \text{Provisions}}{\text{Gross Advances} - \text{Provisions}} \)

iv. Net NPAs / Total Assets

v. Provisions for loan losses / Gross Advances

E. Liquidity

The liquidity or managerial ratio shows how the bank is able to meet its recurring financial obligations. Higher the ratio indicates the stability and safety of the firm to meet the short term financial obligations.

i. Liquid Assets / Total Assets

ii. \( \text{SLR} = \frac{\text{Excess Reserve + Invest (Govt + Approved securities) + Current A/C Bal}}{\text{Total Demand and Time Liability}} \)

iii. \( \text{CRR} = \frac{\text{Current Account Balances held with RBI}}{\text{Net Demand and Time Liabilities}} \times 100 \)

The above ratios were computed and the interpretation of trend was done by calculating the mean, standard deviation and coefficient of variation for each bank groups and a comparative analysis of pre- and post-reform period were also presented.

Stationarity Analysis:

The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (1976 and 1979). The basic objective of the test is to examine the null hypothesis that \( \phi = 1 \), against the one-sided alternative \( \phi < 1 \). Thus the hypotheses framed as:

\( H_0: \text{Series contains a unit root} \)

\( H_1: \text{Series is stationary.} \)

Dickey-Fuller (DF) can be conducted allowing for an intercept, or an intercept and deterministic trend, or neither, in the test regression. In practice, the following regression was employed for ease of computation and interpretation so that a test of \( \phi = 1 \) is equivalent to a test of \( \psi = 0 \) (since \( \phi - 1 = \psi \)).

\[ \Delta y_t = \psi y_{t-1} + \mu_t + \lambda_t + u_t \]  

(1)

The model for the unit root test in each case is

\[ y_t = \phi y_{t-1} + \mu + \lambda_t + u_t. \]  

(2)

Which can be written as:

\[ \Delta y_t = \psi y_{t-1} + \mu + \lambda_t + u_t \]  

(3)

The test statistics for the DF tests are defined as \( \text{test statistic} = \frac{(\psi^\prime)/(SE(\psi^\prime))}{\psi^\prime}. \) The null hypothesis of a unit root is rejected in favour of the stationary alternative in each case if the test statistic is more negative than the critical value.
The simple Dickey-Fuller unit root test described above is valid only if the series is an AR (1) process. If the series is correlated at higher order lags, the assumption of white noise disturbances $u_t$ is violated. The augmented Dickey-Fuller (ADF) test constructs a parametric correction for higher-order correlation by assuming that the $y$ series follows an AR ($p$) process and adding $p$ lagged difference terms of the dependent variable $y$ to the right-hand side of the test regression. The solution is to ‘augment’ the test using $p$ lags of the dependent variable. The alternative model is written as:

$$ \Delta y_t = \psi y_{t-1} + \sum_{i=1}^{n} \alpha_i \Delta y_{t-i} + u_t \quad (4) $$

The lags of $\Delta y_t$ now ‘soak up’ any dynamic structure present in the dependent variable, to ensure that $u_t$ is not auto-correlated. The test is known as an augmented Dickey- Fuller (ADF) test and is conducted on $\psi$, and the same critical values from the DF tables are used as before. The augmented Dickey-Fuller (ADF) test statistics are calculated independently for each variable, with appropriate lags to adjust for autocorrelation. The adjusted test statistics, [adjusted using the tables in Im, Pesaran, and Shin, 1995] are distributed as N (0, 1) under the null of a unit root and large negative values lead to the rejection of a unit root in favor of stationarity.

**Trend analysis:**

To find the magnitude and the direction of change in the variables during the study period for different banks, growth rates were estimated. As the performance of the commercial banks normally grow from year to year and each year enables it to have an enlarged base to compound the growth rate, hence exponential growth function was fitted to analyze the trends in selected parameters. The equation of the exponential curve is of the form:

$$ Y = Ab^x e^u $$

where $Y$ is the variable, $x$ years, $A$ is the $Y$ intercept, $b$ slope of the function and $u$ random factors. Under the growth function, the growth rate ($G$) in percentage form shall be:

$$ G = [ \text{Antilog of Log } b - 1 ] \times 100. $$

Apart from the exponential growth rate, percentage growth rate over the base year was calculated to analyze the trends on year-to-year basis. The percentage growth rate over the base year is given as:

$$ g = \frac{(V_c - V_p)}{V_p} \times 100 $$

where: $g$ = annual rate of growth.

$V_c$ = Value of the given parameter in the current year.
\( V_p = \text{Value of the given parameter in the previous year.} \)

**Variable description**

Berger and Humphrey (1997) pointed out that, although there is no ‘perfect approach’, the intermediation approach may be more appropriate for evaluating entire financial institutions because this approach is inclusive of interest expenses, which often account for one-half to two-thirds of total costs. Moreover, the intermediation approach may be superior for evaluating the importance of frontier efficiency of financial institutions. The main consequence of the intermediation approach is that deposits are considered as inputs, and interest on deposits is a component of total costs, together with labour and capital costs.

**TABLE-3.1**

**VARIABLES USED FOR ESTIMATION**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Terminology</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>Total loan</td>
<td>Loans issued</td>
</tr>
<tr>
<td>Output</td>
<td>Other earning assets</td>
<td>Sum of investment securities</td>
</tr>
<tr>
<td>Input</td>
<td>Labour</td>
<td>Number of full-time employees</td>
</tr>
<tr>
<td>Input</td>
<td>Fixed capital</td>
<td>Fixed Assets (Property and Equipment)</td>
</tr>
<tr>
<td>Input</td>
<td>Customer and short-term funding funds.</td>
<td>Deposits and current accounts</td>
</tr>
<tr>
<td>Input price</td>
<td>Labour</td>
<td>Total staff costs / Number of employees</td>
</tr>
<tr>
<td>Input price</td>
<td>Fixed capital</td>
<td>Exps on fixed assets/Book value of fixed assets</td>
</tr>
<tr>
<td>Input price</td>
<td>Customer and short-term funding funds.</td>
<td>Interest expense/Deposits</td>
</tr>
</tbody>
</table>

Accordingly, this study specifies two outputs: total loans and other earning assets and three inputs: labour, fixed capital and total customers and short term funding and the prices are personnel expenses to average number of personnel for labour, total capital expenses to total fixed assets for fixed capital and interest expenses to average total customers and short term funding for total customers and short term funding. This assumption on outputs is consistent with studies conducted by Fuentes and Vegara (2003), Kraft *et al.*, (2002) and Vivas (1997). Other assets as a third output, was excluded in this study as at least two of the banks did not have any values recorded for these variables. Since the model variables are in logarithms, this would have introduced an error in the inputs, hence the exclusion. Table 3.1 outlines the data extracted from the bank’s Annual Reports that were used in the estimation of frontiers.

**DEA Modeling**
DEA is a linear programming technique initially developed by Charnes, Cooper and Rhodes (1978) to evaluate the efficiency of public sector non-profit organizations, based on earlier work initiated by Farell (1957). It was later extended by Banker et al. (1984). Sherman and Gold (1985) were the first to apply DEA to banking. DEA calculates the relative efficiency scores of various Decision-Making Units (DMUs) in the particular sample. The DMUs could be banks or branches of banks. The DEA measure compares each of the banks to branches in that sample with the best practice in the sample. It tells the user which of the DMUs in the sample are efficient and which are not. The ability of the DEA to identify possible peers or role models as well as simple efficiency scores gives it an edge over other methods. As an efficient frontier technique, DEA identifies the inefficiency in a particular DMU by comparing it to similar DMUs regarded as efficient, rather than trying to associate a DMU’s performance with statistical averages that may not be applicable to that DMU.

Fa’re et al. (1994) also propose an ‘enhanced decomposition’, which takes the efficiency change component calculated relative to the constant returns to scale (CRS) technology and further decomposes into a ‘pure technical efficiency change’ component (calculated relative to the variable returns to scale (VRS) technology) and a residual ‘scale efficiency’ component, which captures changes in the deviation between the VRS and CRS technologies. The decomposition becomes,

$$M_0(y_s, x_s, y_t, x_t) = TC \times TE \times SE$$

where TC represents Technical Change, TE represents pure efficiency change and SE represents scale change. The scale change and pure efficiency change components are the decomposition of the efficiency component $TEC = TE \times SE$. DEA modeling allows the analyst to select inputs and outputs in accordance with managerial focus. This is an advantage of DEA since it opens the door to what-if analysis. Furthermore, the technique works with variables of different units without the need for standardisation (e.g. dollars, number of transactions, or number of staff). The approach used in this study was variation of the intermediation approach. The input variables used in this study are labour, fixed capital, customer and short-term funding funds and the output variables are total loans and other earning assets. Appropriate GDP deflators have been used to deflate the values.

**Constant Returns to Scale (CRS or CCR):**

Charnes, Cooper and Rhodes (1978) proposed a model which had input orientation and assumed constant return to scale (CRS). Assume that there are $n$ DMUs consuming varying amounts of $k$ different inputs to produce $m$ different outputs. Specifically, $DMU_j$ are represented
by \( x_i \) and \( y_i \) respectively. The \( K \times N \) input matrix, \( X \) and \( M \times N \) output matrix \( Y \) represent the data of \( N \) firms. The purpose of DEA was to construct the frontier that the observed data lie on or below the production frontier. For each DMU a measured ratio of all outputs \((y)\) over all inputs \((x)\), such as \( u'y/v'x \), where \( u \) is output weight and \( v \) is input weight. To select optimal weights, the mathematical programming problem:

\[
\text{Max}_{u,v} \left( \frac{u'y}{v'x} \right) \\
\text{s.t} \quad \frac{u'y_j}{v'x_j} \leq 1 \quad j = 1, 2, \ldots, N \\
u, v \geq 0
\]  

(1)

The finding values for \( u \) and \( v \), such that the efficiency measure of the \( i \)th DMU is maximised, subject to the constraint that all efficiency measures must be less than or equal to one. One problem with this particular ratio formulation is that it has an infinite number of solutions. To avoid this, an equivalent envelopment form of this problem is estimated, \( i.e. \)

\[
\text{Min}_{\theta, \lambda} \theta \\
\text{s.t} \quad -y_i + y\lambda \geq 0 \\
\theta x_i - X\lambda \geq 0 \\
\lambda \geq 0
\]  

(2)

The value of \( \theta \) obtained will be the efficiency score of \( i \)th DMU. It will satisfy \( \theta \leq 1 \), with a value of 1 indicating the point on the frontier and technical efficient DMU.

**Super Efficiency:**

Andersen and Petersen's model for estimating super-efficiency score for DMU \( k \) (denoted by \( \text{TE}^{k, \text{super}} \))

\[
\text{TE}^{k, \text{super}} = \theta^{\text{super}}_k - \varepsilon \left[ \sum_{i=1}^{m} s_i^- + \sum_{i=1}^{m} s_r^+ \right]
\]

\[
\text{Min} \theta^{\text{super}}_k, \lambda, s_i^-, s_r^+
\]

Subject to

\[
\sum_{j=1, j\neq k}^{m} \lambda_j y_{rj} - s_r^+ = y_{rk} \quad r = 1, 2 \ldots s
\]

\[
\sum_{j=1, j\neq k}^{n} \lambda_j x_{ij} + s_i^- = \theta^{\text{super}}_k x_{ik} \quad r = 1, 2 \ldots m
\]

\( s_i^-, s_r^+ \geq 0 \)

\( \lambda_j (j \neq k) \geq 0 \)

\( j = 1, 2 \ldots n. \)
The difference between Super-efficiency model and standard efficiency model is that in super models the DMU₀ (the DMU evaluated) is eliminated from the reference set (indicated by j≠0 in the LP). The Super-efficiency score can be greater than 1.

**Variable Returns Scale (VRS or BBC):**

The CRS assumption is appropriate only when the firm is operating under constant returns to scale. Due to certain constraint the firm may be unable to operate at optimal scale. BCC (Banker et al., 1984) made an extension of CRS to VRS situations. This model is modified by convexity constraint \(N₁'λ = 1\) to (2) hence

\[
\begin{align*}
\text{Min } θ, λ \quad θ \\
s.t \quad & -yᵢ + yλ ≥ 0 \\
& θxᵢ - Xλ ≥ 0 \\
& N₁'λ = 1 \\
& λ ≥ 0 
\end{align*}
\]

(3)

\(N₁\) is \(N × 1\) vector of one. This convex hull of intersecting will tighten the data points than in CRS and provide the efficiency score greater than or equal to the value obtained from CRS model.

**Scale Efficiency:**

The score obtained from CRS model and VRS model are calculated and their difference are calculated as scale efficiency of that DMU. This can be expressed by substituting \(N₁'λ = 1\) restriction with \(N₁'λ ≤ 1\) in (3)

\[
\begin{align*}
\text{Min } θ, λ \quad θ \\
s.t \quad & -yᵢ + yλ ≥ 0 \\
& θxᵢ - Xλ ≥ 0 \\
& N₁'λ ≤ 1 \\
& λ ≥ 0 
\end{align*}
\]

(4)

The scale efficiency refers to the firm’s ability to work at the optimal scale since \(TE_{CRS} = TE_{VRS} × SE\). This can be expressed as \(SE = TE_{CRS} / TE_{VRS}\). Corresponding to these efficiency measures, the measures of inefficiency can be obtained as \((TEᵢ⁻¹ - 1)\), \((PTEᵢ⁻¹ - 1)\) and \((SEᵢ⁻¹ - 1)\), respectively (Isik and Hassan, 2002).

**Cost and Allocative Efficiency:**

Depending on whether one is interested in maximizing outputs or minimizing inputs, one can calculate the input-oriented technical efficiency which shows how much a company’s inputs should
be decreased to be efficient leaving outputs unchanged, or output-oriented technical efficiency which presents how much a company’s productivity should be increased using the same values of inputs. The input-oriented analysis is particularly useful for evaluating banks’ performance as it measures cost efficiency. The VRS model of DEA model was used for cost minimization. The study used the standard DEA approach; the efficiency measures are the result of the implementation of a variable returns to scale (VRS), input-oriented cost minimization model (eq 2). Since price information was available, both technical (TE) and allocative efficiency (AE) were measured. Allocative efficiency in input selection involves selecting that mix of inputs which produce a given quantity of output at a minimum cost, given the prevalent input prices. Allocative and technical efficiency combine to provide an overall measure of economic efficiency.

\[
\text{Min } \lambda, x_i \quad w_i'x_i^* \\
\text{s.t} \quad -y_i + y \lambda \geq 0 \\
\quad x_i^* - X \lambda \geq 0 \\
N1^*\lambda = 1 \\
\lambda \geq 0
\]  

Thus, the cost efficiency of any given bank is obtained as the ratio of minimum cost to the observed cost written as follows: \( CE = w_i'x_i^*/w_i'x_i \). The total cost efficiency, or economic efficiency, of the ith observation was calculated as the ratio of minimum cost to observed cost. Cost efficiency (CE) can be seen also as the product of technical and allocative efficiencies (\( CE = TE \times AE \)). In other words, firms having higher costs than the frontier may be so either because they are not using the most efficient technology (technical inefficiency) or/and because they are not using the cost minimizing input mix (allocative inefficiency). The allocative efficiency can be derived by \( AE = CE/TE \). The long run sustainability of any banking sector depends on the economic efficiency. The banks is economically efficient if it was both technical and price efficiency. Corresponding to these efficiency measures, the measures of inefficient can be obtained as \((CE_i^{-1} - 1), (TE_i^{-1} - 1)\) and \((AE_i^{-1} - 1)\), respectively (Isik and Hassan, 2002; Welzel and Lang, 1997).

**Derivation of productivity index:**

The Malmquist Productivity Index (MPI) introduced by Caves, Christensen and Diewert (1982) is a normative measure that constructs a production frontier representing the technology and uses the corresponding distance functions evaluated at different input-output combinations for productivity comparing the proportion of output to input over two periods of time. In order to estimate efficiency and productivity changes over time, the Malmquist Productivity Index (MPI) was calculated. The Malmquist TFP index measures the TFP change between two data points...
by calculating the ratio of the distances of each data point relative to a common technology. According to Fare et al. (1994) the Malmquist (output oriented) TFP change index between periods, the base period and period \( t \) is given by:

\[
M_0(\text{Y}_s, \text{X}_s, \text{Y}_t, \text{X}_t) = \left( \frac{d_s^0(\text{Y}_t, \text{X}_t)}{d_s^0(\text{Y}_s, \text{X}_s)} \times \frac{d_0^s(\text{Y}_t, \text{X}_t)}{d_0^s(\text{Y}_s, \text{X}_s)} \right)^{1/2}
\]

where the notation \( d_s^0(\text{Y}_t, \text{X}_t) \) represents the distance from the period \( t \) observation to the period’s technology. A value of \( M_0 \) greater than one will indicate positive TFP growth from the period \( s \) to period \( t \) while a value less than one indicates TFP decline.

**Total Factor Productivity Change Index**

\[
\text{Total Factor Productivity Change Index} = \frac{D'(y', x')}{D'^{t+1}(y'^{t+1}, x'^{t+1})} \times \left[ \frac{D'^{t+1}(y_{t+1}, x_{t+1})}{D'(y', x')} \times \frac{D'^{t+1}(y', x')}{D'(y', x')} \right]^{1/2}
\]

**Technological Change Index**

\[
\text{Technological Change Index} = \left[ \frac{D'^{t+1}(y_{t+1}, x_{t+1})}{D'(y'^{t+1}, x'^{t+1})} \times \frac{D'^{t+1}(y', x')}{D'(y', x')} \right]^{1/2}
\]

**Technical Efficiency Change Index**

\[
\text{Technical Efficiency Change Index} = \left[ \frac{D'^{t+1}(\text{CRS})(y_{t+1}, x_{t+1})}{D'(\text{CRS})(y', x')} \right]
\]

**Pure Technical Efficiency Change Index**

\[
\text{Pure Technical Efficiency Change Index} = \left[ \frac{D'^{t+1}(\text{VRS})(y_{t+1}, x_{t+1})}{D'(\text{VRS})(y', x')} \right]
\]

**Scale Efficiency Change Index**

\[
\text{Scale Efficiency Change Index} = \left[ \frac{D'^{t+1}(\text{CRS})(y_{t+1}, x_{t+1})}{D'(\text{CRS})(y', x')} \right] \times \left[ \frac{D'^{t+1}(\text{VRS})(y_{t+1}, x_{t+1})}{D'(\text{VRS})(y', x')} \right]^{-1}
\]

Fa¨re et al. (1994) also recommended an ‘enhanced decomposition’, which takes the efficiency change component calculated relative to the constant returns to scale (CRS) technology and further decomposes into a ‘pure technical efficiency change’ component (calculated relative to the variable returns to scale (VRS) technology) and a residual ‘scale efficiency’ component, which captures changes in the deviation between the VRS and CRS technologies. The decomposition becomes

\[
M_0(\text{Y}_s, \text{X}_s, \text{Y}_t, \text{X}_t) = \text{TC} \times \text{TE} \times \text{SE}
\]

Where TC represents technical change, TE represents pure efficiency change and SE represents scale change. The scale change and pure efficiency change components are the decomposition of the efficiency component \( \text{TEC} = \text{TE} \times \text{SE} \).

**Isotonic Regression:**
To ensure the validity of the DEA model specification, an isotonicity test was done. In numerical analysis, isotonic regression (IR) involves finding a weighted least-squares fit $x \in \mathbb{R}^n$ to a vector $a \in \mathbb{R}^n$ with weights vector $w \in \mathbb{R}^n$ subject to a set of non-contradictory constraints of kind $x_i \geq x_j$. Such constraints define partial order or total order and can be represented as a directed graph $G = (N, E)$, where $N$ is the set of variables involved, and $E$ is the set of pairs $(i, j)$ for each constraint $x_i \geq x_j$. Thus, the IR problem corresponds to the following quadratic program (QP):

$$\begin{align*}
\text{Min} & \sum_{i=1}^{n} w^i (x^i - a^i)^2 \\
\text{s.t:} & \quad x_i \geq x_j \quad \text{for all } (i, j) \in E.
\end{align*}$$

**Correlation Analysis:**

To identify the inter-relationship between the efficiency score and the basic indicators of profitability, productivity, operation, liquidity and asset quality, the correlation analysis was done to explain variation in estimated efficiency scores to explanatory variables. It should be noted that dependent variable are the efficiency score which was derived from the DEA analysis. The results will be based on the sign of the co-efficient, ie, a positive coefficient depicts that the variable is positively related to efficiency while a negative coefficient means the variable causes a decline in efficiency. Spearman correlation coefficients are used, which also indicates which hypothesized relationships are supported by the analysis.

**Regression Analysis:**

To identify which factors influences the resulted DEA efficiency score, the scores were regressed to study impact of various market and regulatory initiatives on efficiency and profitability of banks. The results will be on the basis of the sign of the co-efficient ie, a positive coefficient depicts that the there is an increase in the efficiency while a negative coefficient indicates there is a decline in its efficiency. As a further robustness check, the study report the results based on step-wise regressions, which eliminate the variables with ‘trivial’ explanatory power.

**Determinant of Profitability Efficiency Model**

Multiple regression analysis has been used to build the model to establish the relationship between profitability and the influencing factors of profitability group-wise. The functional form of regression model is:

$$Y = \beta_0 + \beta_1 \text{CE} + \beta_2 \text{PR} + \beta_3 \text{SIZE} + \beta_4 \text{HHI} + \varepsilon$$
The dependent variable Y represents gross profit margin as percent of total assets (ROA) and independent variables are cost efficiency scores (CE), priority sector advances to total advances (PR), logarithm of total assets (SIZE), Herfindahl-Hirschman “concentration” index (HHI), $\beta_0$ Intercept, $\beta_i$ (i=1,2,3,4) coefficients of regression parameters and $\epsilon$ random term. 

**Concentration Index:**

In order to judge the performance each bank groups, it becomes imperative to know the relative efficiency of each group. For this purpose, Herfindhal’s index of concentration has been computed. Herfindhal’s index of concentration has been defined as:

$$H_i = \sum_{i=1}^{n} \frac{V_i^2}{\sum V_i}$$

where $H_i$ = Overall deposit.

$V_i$ = ’i’th unit’s share of deposit

n = Number of units

**Impact of Market Regulation**

To examine the impact of market regulations on the efficiency of banking systems, the following model was estimated:

$$TE = \beta_0 + \beta_1 FD + \beta_2 PI + \beta_3 FOG + \beta_4 T_1 + \beta_5 T_2 + \beta_6 T_3 + \beta_7 T_4 + \epsilon$$

Here the dependent variable is technical efficiency from DEA model (TE) and the independent variables are fiscal deposit as a percentage to GDP (FD), private investment (PI), share of foreign banks in total credit (FOG), dummy for periodisation 1990 ($T_1$), dummy for periodisation 1994 ($T_2$), dummy for periodisation 1999 ($T_3$), dummy for periodisation 2006 ($T_4$), $\beta_0=\text{Intercept}$, $\beta_i=\text{Coefficients of regression parameters}$ ($i=1,2,3,4$), and $\epsilon=\text{random term}$. 

**Periodisation:**

According to Hansen (2001) and Perron (2005) changing of dependent variable at dates in the sample period plays an empirically relevant role in applied time series analysis. There has been a large volume of work targeted at developing testing and estimation methodologies for regression models which allow for change. They are estimated using the technique developed by Bai (1997) and Bai and Perron (1998). 

The standard multiple linear regression models start with T periods and m potential breaks (producing m+1 regime). For the observations $T_i$, $T_{i+1}$...T_{i+1-1} in regime j the regression model is:

$$Y_t = X_t'\beta + Z_t'\delta_j + \epsilon_t$$  \hspace{1cm} (6)
for the regimes $j=0,\ldots,m$. The regressors are divided into two groups. The $X$ variables are those whose parameters do not vary across regimes, while the $Z$ variables have coefficients that are regime-specific. While it is slightly more convenient to define break dates to be the last date of a regime, the present study follow EViews’s convention in defining the break date to be the first date of the subsequent regime and tie down the endpoints by setting $T_0=1$ and $T_{m+1}=T+1$. Once the number and identity of the breakpoints is determined, the model may be estimated using standard regression techniques. Thus, this may rewrite the equation specification as a standard regression equation

$$Y_t = X_t'\beta + Z_t'\delta_j + \epsilon_t$$

with fixed parameter vectors $\beta$ and where $\delta = (\delta_0', \delta_1', \ldots, \delta_m')$ is an expanded set of regressors interacting with the set of dummy variables corresponding to each of the $m+1$ regime segments. The breakpoints may be known apriori or they be estimated using a variety of approaches. The breakpoint estimation methods may broadly be divided into two categories: global maximizers for the breakpoints and sequentially determined breakpoints.

In matrix notation, the standard regression may be written as:

$$y = X\beta + \varepsilon$$

where $y$ is a $T$ – dimensional vector containing observations on the dependent variable, $X$ is a $T \times K$ matrix of independent variables, $\beta$ is a $k$–vector of coefficients, and $\varepsilon$ is a $T$–vector of disturbances. $T$ is the number of observations and $X$ is the number of right-hand side regressors.

All the above calculations were done using EViews, DEAP, EMS and SPSS software packages.

**Limitation of the study**

The study was based on the annual data of the banks. If quarterly or monthly observation of the variable had been available, the number of observation could have been increased and the reliability of estimation could have improved. The quality of the research also depends on the quality and reliability of the data published in Annual Report published in RBI. The study was based on the aggregate level data. If it had been narrowed to the regional or branch level, more of the variation could be captured. The study was based on common size available in the annual report hence it is subject to its limitation of common standard size. This study concentrated on the business strategy concept of evaluating the efficiency and the performance of banking sector and eliminated the customer service as non-value added
strategy. Among the four methods available for estimating the frontier the study had used only DEA method and had not tried other methods due to time constraint.

For certain indicators the data were available only from 1990 onwards for specific groups or for all groups. Similarly the cost of input relating to foreign and private sector banks were available from 1991 onwards only. Hence, the analysis portraits to the periods from which the data was available.

Despite these limitations, the findings do not negate the conclusions draw and will definitely add to the existing literature and provides scope for further research in future.