CHAPTER - 3
DATA BASE AND METHODOLOGY

The importance of any empirical study is generally examined and valued by its data base and methodology. Data base generally includes decision on four areas; these are: (i) the choice of the field, (ii) definition and recognition of the variables / parameters on which the data are to be collected, (iii) methods and sources of data collection and (iv) problems and limitations, if any of the collected data. The Indian Scheduled Commercial Banks have been chosen as the field of our study and the period of our analysis is the post-reforms era ranging from 1992 to 2007. Other points related to data base have been described here on the basis of those parameters which are useful for knowing and analysing sources and uses of banking funds. The selection of methods for analysing the data sets is another crucial part of the study. According to the objectives and hypotheses of our study, we have chosen different methods, which have been briefly explained here.

3.1 Data Base for Growth Measures

The basic data that have been collected for the purpose of analyzing the growth rates in different components of bank funds are on sources of funds, like, total deposits, borrowings, other liabilities and provisions. We have collected data on the uses of funds too, like data on cash in hand and balance with RBI, other liquid assets, advances and on investments and their components. We have also collected data on the performance parameters such as data on provision and contingencies, profit and loss, interest expended, payments to employees, other operating expenses and their components. Besides data on non-performing assets, data on capital adequacy ratio have also been collected. We have collected data on whole sale price index number by which the respective data have been deflated to get the real picture of banking sector in India for the period of our study. The data have been collected for the Scheduled Commercial Banks as a whole as well as for other broad groups, like, Nationalized Banks, State Bank of India and its Associates, Private Banks and Foreign Banks. Unlike other nationalized banks the State Bank of India and its Associates (SBI and its Asso) possess a different character and status. We have also collected these data
for four public sector banks, three private banks and for three foreign banks individually as samples banks for the purpose of our study. Actually, the number of banks in each category has been determined on the basis of proportionate stratified random sampling method. Thus, the banks selected are State Bank of India (SBI), Allahabad Bank, Punjab National Bank (PNB) and United Commercial Bank (UCO) as sample banks among the public sector banks category. In the private banks category the banks selected as samples for the study include ICICI Bank, Bank of Rajasthan and Dhanalakshmi Bank. Among the foreign banks, the sample foreign banks selected for the purpose of our study are ABN Amro Bank, Hong Kong & Shanghai Bank (HSBC) and Standard Chartered Bank.

The sources of these data are mainly (i) Statistical Tables Relating to Banks in India. (ii) Handbook of Statistics on Indian Economy, RBI. In addition to these, data on whole sale price index for all commodities combined has been collected to exclude the effect of inflation, from (a) Handbook of Statistics on Indian Economy, RBI, (Various issues) and (b) Statistical Outline of India.

3.2 Data Base for Ratio Analysis

Ratio analysis for banking business slightly differs from usual ratio analysis. For the analysis on liquidity ratios of the banks, the necessary data which have been collected are on current assets current liabilities, liquid assets, liquid liabilities, quick assets etc. Further, for determining the profitability ratios the collected data are on interest spent on deposits and borrowings, interest income, operating expenses, miscellaneous income, provisions made during the year and average working funds. For calculating the efficiency ratios we have collected data on actual interest paid on various deposits, actual interest paid on borrowings from various sources, actual interest earned on loans and advances, average deposits, average loans and advances, total investments etc. All these data have been collected for each of the individual sample banks as well as for the Nationalized Banks, State Bank of India and its Associates, Private Banks and Foreign Banks as a whole for a period of 16 years i.e. from 1992 to 2007. The sources of these data are (i) Statistical Tables Relating to Banks in India (Various annual publications), RBI, (ii) Report on Trend and progress of Banking in India (various annual issues) and (iii) the website http://www.rbi.org.in.
3.3 Data Base for Banking Efficiency Assessment

Apart from ratio analysis, to assess the efficiency of the banking business, we have used Data Envelopment Analysis, the detailed methodology of which has been provided in the methodology section of our study. The requisite data for this analysis are on deposits, borrowings, payment to employees, other operating expenses, total advances, investments and net interest margin.

We have collected these data for the scheduled Commercial Banks as a whole, as well as for the Nationalized Banks, State Bank of India and its Associates, Private Banks and Foreign Banks. These data have been collected for the sample banks as well, for a period of 16 years i.e. from the year 1992 to 2007.

The sources of these data are (i) Statistical Tables Relating to Banks in India (Various annual publications) and (ii) the website http://www.rbi.org.in.

It should be mentioned here that the study is solely based on secondary sources of data and due to non-availability of data, in some cases the study has to be restricted as per the availability of data on the banks. Some of the banks like Bank of Madura, New Bank of India etc. have been merged with other banks. Therefore, their aggregated data for the specified period of 16 years are not available. There are still other banks like American Express Bank, Bank of Tokyo, Bank of America, etc. whose data for the initial years are not available. A few new generation private banks have crept up in the recent past; naturally their data for the specified period of 16 years are not available. Here, in this study only those banks have been taken into consideration for random sampling, whose data for at least twelve years were available. In other words, although we have used stratified random sampling technique for the sample selection of banks, if a bank whose data for the specified number of years were not available, it has been excluded from the sample selection.

3.4 Methodology of Growth Measures

It is seen from the review of past literature that most of the researchers have concentrated on measuring performance of the different groups of banks in India and that too giving stress on profit as the main criterion for performance analysis of the banks. But as per the objectives of our study, we have analysed the sources and uses
of bank funds and their components along with profit in order to judge the performance of the scheduled commercial banks over time.

Apart from the computation of averages, standard deviation, ratios and proportions, the performances of the banks have been examined by estimating trend growth rates of different performance indicators relating to bank-funds viz. assets, deposits, investments, loans and advances, income-expenditure components, profit, etc. For the estimation of trend growth rate, kinked exponential trend equation \( \ln y_t = a + b_1 D_{1t} + b_2 D_{2t} + u_t \) has been fitted and the estimated values of \( b_1 \) and \( b_2 \) have been taken as the trend growth rates of first and second sub-periods respectively. Here \( t = -7.5 \) (for 1992), -6.5 (for 1993), ........... , 6.5 (for 2006), 7.5 (for 2007), and \( D_j = 1 \) for the jth sub-period and 0 otherwise (where \( j = 1, 2 \)). Further, \( u_t \) is the disturbance term satisfying the usual assumption of the classical linear regression model. To test the statistical significance of each trend growth rate, its standard error has been calculated and 't' test has been applied. The 't' values or standard errors have also been reported to indicate the probability levels at which the coefficients are statistically significant.

The statistical criterion that has been used for selection of the trend equation is \( R^2 \) (Adjusted \( R^2 \)) which is a measure of goodness of fit. The F values and DW statistics have been reported to indicate respectively the levels of significance of \( R^2 \) and the presence or absence of autocorrelation problem in the disturbance term [Boyce (1987), Pp. 267-271].

### 3.5 Methodology of Ratio Analysis

As banking industry has grown in size, outreach and complexity, there has been an increasing concern as well as interest regarding the examination of financial viability, accountability and efficient management of human and financial resources in the banking sector. Various tools might be used to evaluate those issues. But, perhaps, ratio analysis has been accepted widely as one of the most favourable tools for the purpose.

We have selected broadly three categories of ratios for the purpose of our study namely (1) Liquidity ratios (2) Profitability ratios and (3) Efficiency ratios.

(1) **The liquidity ratios that have been used in the study are:** (i) Current ratio (ii) Liquid ratio and (iii) Absolute liquid ratio.
(i) Current ratio is the ratio of current assets to current liabilities. Current assets include cash and balance with RBI, balance with banks, money at call and short notice, advances other than term loans and other assets. Current liabilities are, on the other hand, consisted of demand deposits, savings bank deposits, other liabilities and provisions.

(ii) Liquid ratio is the ratio of liquid asset to liquid liability. Liquid asset is the summation of cash and balance with RBI, balance with banks, money at call and short notice and advances other than term loans. Liquid liability includes only demand deposits and savings bank deposits.

(iii) Absolute liquid ratio is the ratio of quick assets to liquid liabilities. Quick assets imply cash, balance with RBI and money at call and short notice.

(2) The ratios that have been computed in the study relating to banking profitability are: (i) Financial return (ii) Financial cost (iii) Operating expenses (iv) Miscellaneous income and (v) Risk cost. All of these ratios are expressed in percentage form.

(i) Financial return = \[
\frac{\text{Interest income}}{\text{Average working fund}} \times 100
\]

Where, average working fund is the sum total of all assets net of contra items.

(ii) Financial cost = \[
\frac{\text{Interest spent on deposits & borrowings}}{\text{Average working fund}} \times 100
\]

(iii) Operating expenses = \[
\frac{\text{Operating expenses}}{\text{Average working fund}} \times 100
\]

(iv) Miscellaneous Income = \[
\frac{\text{Miscellaneous Income}}{\text{Average working fund}} \times 100
\]

(v) Risk Cost = \[
\frac{\text{Provisions made during the year}}{\text{Average working fund}} \times 100
\]
(3) Efficiency ratios that have been considered in our study are as follows:

(i) Interest cost of deposits = \[
\frac{\text{Actual interest paid on various deposits}}{\text{Average deposits}} \times 100
\]

Where, average deposits do not include demand deposits.

(ii) Interest cost of borrowings = \[
\frac{\text{Actual interest paid on borrowings from various sources}}{\text{Average borrowings outstanding}} \times 100
\]

(iii) Interest yield on loans = \[
\frac{\text{Actual interest earned on loans & advances}}{\text{Average loans & advances}} \times 100
\]

(iv) Interest yield on investment and Bank balances = \[
\frac{\text{Actual interest earned on investment and bank balances}}{\text{Average bank balances and investment}} \times 100
\]

(v) CD Ratio (Credit to Deposit ratio) = \[
\frac{\text{Total loans & advances outstanding}}{\text{Total deposits outstanding}} \times 100
\]

These ratios have been calculated for the whole period (taking average figures of 16 years) and for the two sub-periods (taking average figures of 8 years for each sub-period); 1st sub-period has been considered from 1992 to 1999 which reflects the initial state of the post-liberalization era and the 2nd sub-period from 2000 to 2007 which reflects the matured state of the post-liberalization era. Both cross-sectional (over the banks) and time series analysis (for each bank over different periods) have been made. It is necessary to mention here, that among the profitability ratios, the banker’s motive must be to maximize the financial return and miscellaneous income and to minimize the financial cost, operating cost and risk cost. Hence the profitability ratios have been categorized into two: profitability-return (financial return and miscellaneous income) and profitability-expense (financial cost, operating cost and risk cost) ratios.
Similarly, among the efficiency ratios, the banker's aim should be to maximize the yields and minimize the costs. Therefore, the efficiency ratios have also been grouped into two: efficiency yield (interest yield on loans, interest yield on bank balances and investment and credit-deposit ratio) and efficiency-cost (interest cost of deposit and interest cost of borrowing) ratios.

For each of these ratios, geometric mean has been considered for calculating average value and co-efficient of variation has been computed for obtaining relative measure of dispersion among banks within a group and over the years for each bank.

3.6 Methodology on Measurement of Banking Efficiency through DEA

DEA [Data Envelopment Analysis] has been one of the widely used measures of efficiency, in the banking sector as it allows comparison of relative efficiency of individual banks and also peer group performances. [Debasish (2006)]. ‘Data Envelopment Analysis is a non-parametric mathematical programming technique used for assessing / evaluating and comparing the relative performances of economic units, with minimal prior assumption on input-output relation. The methodology on DEA was originally developed by Charnes, Cooper and Rhodes (1978)’. [Chatterjee and Sinha (2006), P.161]. The economic units are called Decision Making units (DMUs). In banking sector, each bank is a DMU. The resources used by the units are called the inputs. A unit or a DMU converts the inputs into output. ‘From the set of available data, DEA identifies (a) reference points (relatively efficient DMUs) that define the efficient frontier (as the best practice production technology) and evaluate the inefficiency of others (b) interior points (relatively inefficient DMUs) that are below the frontier. Efficiency is equal to ratio of total sum of weighted outputs to total sum of weighted inputs.

\[
\text{Efficiency} = \frac{\text{Weighted sum of Outputs}}{\text{Weighted sum of Inputs}}
\]

One of the basic choices in selecting a DEA model is whether to use input orientation or an output orientation’. [Debasish (2006), P. 327]. In our study we have used the CRR model (i.e., Charnes, Cooper and Rhodes model) which is an output oriented model. The DMUs produce the highest possible amount of output with the given amount of input.
The CRR model measures efficiency of each DMU as the ratio of weighted outputs to weighted inputs. “The weights for the ratio are determined by a restriction that similar ratios for every DMU have to be less than or equal to unity, thus reducing multiple inputs and multiple outputs to single ‘virtual’ input and single ‘virtual’ output without requiring pre-assigned weights. The efficiency measure is then a function of weights of the ‘virtual’ input-output combination. Formally the efficiency measure for the DMU can be calculated by solving the following mathematical programming problem:

Max. \( h_0(u,v) = \{ \sum_{r=1}^{S} U_r Y_r^0 \} / \{ \sum_{i=1}^{m} V_i X_i^0 \} - \) \hspace{1cm} \text{Equation (i)}

\text{Subject to} \{ \sum_{r=1}^{S} U_r Y_r \} / \{ \sum_{i=1}^{m} V_i X_i \} \leq 1, j=1,2, \ldots, n - \text{Equation (ii)}.

\( U_r \geq 0, r = 1,2, \ldots, s. \)

\( V_i \geq 0, i = 1,2, \ldots, m. \)

Where \( x_{ij} = \text{observed amount of input of the } i^{th} \text{ type of the } j^{th} \text{ DMU} (X_{ij} > 0, i = 1,2, \ldots, m, j = 1,2, \ldots, n) \) and \( y_{rj} = \text{observed amount of the } i^{th} \text{ type of the } j^{th} \text{ DMU} (y_{rj} > 0, r = 1,2, \ldots, s, j = 1,2, \ldots, n). \) The variables \( U_r \) and \( V_i \) are the weights to be determined by the above programming problem. ‘s’ is the total number of input variables and ‘m’ is the total number of output variables.”

[Debasish (2006), P. 327-28.]

In our study, we have used total investment, total advances, net interest margin (difference between interest earned and interest expended) as the output variables and total deposits and payment to employees as the input variables. So output variables reflect mainly uses of funds and in input variables there is a component, namely total deposits that is a major part of sources of funds. We have calculated the technical efficiency of the 15 banks by assuming the presence of constant returns to scale (as assumption of variable returns to scale is discarded by many existing studies). Further, individual sample banks are taken into consideration along with bank groups in DEA to get the figures of DMUs as 15 which is more than twice the figure of total output and input variables (i.e., 5 in number) because in a study it is pointed out
that “usually, the total number of DMUs should be at least twice the number of inputs plus output factors.” [Debasish (2006), P.327]. “Technical efficiency means the ability to avoid waste by producing as much output as input usage allows, or by using as little input as output production allows”. [Chatterjee and Sinha (2008), P. 25]. We have calculated the technical efficiencies of all the 15 banks of our study during each of the years 1992 to 2007. Banks having technical efficiency exactly 100% are taken as efficient banks and rest are in efficient. However, in our study those banks are categorized as efficient banks whose technical efficiency score is 100%. If efficiency score varies from 80% to less than 100%, the bank concerned is taken as moderately efficient. We have identified a bank as inefficient if its technical score is less than 80%. To make year-wise comparison of technical efficiency scores we have to assume that no major technological upgradation has taken place in the Indian banking sector. But this assumption is not always tenable. To have a meaningful exercise in this regard with a realistic assumption, we have carried out this DEA for banks sub-period wise on the consideration that within a sub-period as defined in the study no major technological break through has taken place in India.

Lastly, with the help of pictorial analysis, the probable causes of banking inefficiencies have been identified and those will be discussed in details in the relevant chapter i.e. Chapter-6.