Chapter-3

RESEARCH METHODOLOGY

This chapter discusses the need, scope, limitations and objectives of the current study. It also lays down the design of the present research work. Further, the statistical tools and techniques used for data collection and its analysis have also been explained to obtain the results.

3.1 NEED FOR THE STUDY

In the modern economy, bankers are to be considered not merely dealers in money but leaders in development. Lending is one of the two principal functions of commercial banks not only because of social obligation to cater to the credit needs of different sections of the community but also because lending is the most profitable activity. By advancing loans, they increase the purchasing power in the economy and influence the aggregate demand and create additional demand in various sectors. Banks through lending contribute to mass production, mass distribution and mass consumption. Following the reforms introduced in the banking sector, the banks have been focusing on increasing their credit portfolios and also identifying new ways to offer credit to all sectors. In changing scenario, there is a need to study the pattern of bank credit. The existing literature reveals that various studies have been conducted on total and sector specific bank credit but studies focusing upon total bank credit by various bank groups are limited. So, the present study is an attempt to examine the credit by various bank groups in post-reforms era.

3.2 OBJECTIVES OF THE STUDY

The study is based upon the following objectives:

1. To appraise the policy of RBI relating to bank credit in India.
2. To study and compare the growth and pattern of bank credit by various bank groups.
3. To examine the factors affecting allocation of credit by various bank groups.

4. To assess the perception of borrowers with regard to bank credit.

5. To make various suggestions on the basis of findings of the study.

3.3 CHAPTER SCHEME

The study has been organized into the following chapters:

1. Introduction
2. Review of Literature
3. Research Methodology
4. RBI Policy on Bank Credit
5. Growth and Pattern of Bank Credit in India
6. Factors affecting Bank Credit in India
7. Borrowers’ Perception regarding Bank Credit
8. Summary and Suggestions.

3.4 SCOPE OF THE STUDY

The present study aims to discuss the structure of credit deployment by Indian commercial banks during the post-reforms era. The study attempts to examine and compare the various aspects of credit deployment by Scheduled Commercial Banks in India, and assesses the perception of customers regarding the loan services provided by the banks. The study is related to Scheduled Commercial Banks (SCBs) excluding Regional Rural Banks (RRBs) operating in India. For the purpose of this study, SCBs have been divided into three broad groups, i.e., public sector banks (SBI & its Associates and nationalised banks), private sector banks, and foreign banks. Both primary as well secondary data has been used for the purpose of this study.
3.5 TIME PERIOD OF THE STUDY

Private and foreign banks were allowed to operate with the recommendation of Narsimham Committee-I in the year 1991, but they made their presence after mid-1990s. Hence, the study covers a span of 15 years from 1997-98 to 2011-12.

3.6 SAMPLE AND SOURCES OF SECONDARY DATA

The sample of this study includes data on all-India basis. Research work covers bank group-wise study and includes all scheduled commercial banks except regional rural banks, i.e., public sector banks (SBI & its Associates, and Nationalised banks), private sector banks and foreign banks. The data pertaining to credit deployed by selected bank groups under study has been collected from the various sources such as Annual Reports of Reserve Bank India (RBI), Trend and Progress of Banking in India, Basic Statistical Return of Scheduled Commercial Banks in India, Handbook of Statistics on Indian Economy, Reports on Currency and Finance, Statistical Tables relating to Banks in India, Monetary and Credit Information Review, Occasional papers of RBI, various books, research papers, articles and various websites of commercial banks.

3.7 SAMPLE AND SOURCES OF PRIMARY DATA

The borrowers’ perception with regard to bank credit has been studied through a primary survey. The population of the sample consists of the individual borrowers residing in northern region. For the purpose of this study, a sample of 550 individual borrowers was selected. For this purpose, sampling was done in three stages taking the highest number of offices as the base. At the first stage, out of seven states/UTs in northern region (division as per RBI) five states/UTs have been selected. The selected states/UTs are Rajasthan, Punjab, Haryana, Delhi and Chandigarh. At the second stage, three districts from each state have been selected. Ajmer, Udaipur and Jaipur represent Rajasthan state; Amritsar, Jalandhar
and Ludhiana represent Punjab state, whereas Ambala, Faridabad and Gurgaon represent the state of Haryana; and the two UTs include Delhi and Chandigarh. At the final stage, different banks have been selected, consisting of 3 public sector banks, 2 private sector banks and 1 foreign bank from each district/UT. As many as 50 borrowers each from the selected districts and UTs have been selected. These borrowers represent the selected bank groups. The division of 50 borrowers among public sector banks, private sector banks and foreign banks was in the ratio of 6:3:1, i.e., 30 borrowers from public sector banks, 15 borrowers from private sector banks, and 5 borrowers from foreign banks.

The primary data was collected through a well-structured questionnaire. In order to validate and find out the reliability of questionnaire, a pilot study was conducted on 50 respondents, before processing the actual one. Appropriate modifications in contents and format of the questionnaire were then incorporated in the light of inconsistencies found during the pilot study.

3.8 VARIABLES FOR ANALYSIS

Growth and performance of bank credit by selected bank groups has been analysed on the basis of eleven parameters, i.e., bank group-wise credit, population group-wise credit, region-wise bank credit, occupation-wise bank credit, credit limit range-wise bank credit, tenure-wise bank credit, type-wise bank credit, security-wise bank credit, sector-wise bank credit, population and occupation-wise bank credit, and region and occupation-wise bank credit. These parameters have been selected on the basis of availability of data as well as the previous studies.

There are numerous factors that affect allocation of credit by various bank groups in India. The variables selected for the study affecting deployment of bank credit by various bank groups in India are capital, deposits, borrowings, investments, non-performing assets (NPAs), profits, number of employees and offices.
3.9 Techniques of Data Analysis

For the purpose of analysis of secondary data, various statistical tools have been applied. One-sample Kolmogorov-Smirnov test, Mean, Standard Deviation, Coefficient of Variation, Exponential Growth Rate and Percentage share have been applied for studying the growth and pattern of bank credit deployed by selected bank groups. Further, ANOVA has also been applied to compare and check the significant difference in credit deployment by selected bank groups. Descriptive analysis has been done on selected factors. Further, correlation has been applied to study the association between different factors which affect the credit deployment pattern; and multiple regression analysis has been used to observe the impact of these independent factors, viz. capital, deposits, borrowings, investments, non-performing assets (NPAs), profits, number of employees and offices on dependent factor of credit.

The analysis of primary data collected through a questionnaire has been done using various statistical tools, such as One-sample Kolmogorov-Smirnov test, Simple percentages, Chi-square test, Kruskal-Wallis and Kendall’s coefficient of concordance test. The satisfaction level and agreement level of the respondents was ascertained on a five-point Likert scale. The reliability and validity of the agreement level to the statements has been checked through reliability analysis.

All statistical calculations have been made by the use of Microsoft Excel and Statistical Package for Social Sciences (SPSS) version 20.

A brief description of all the above mentioned statistical tools and formulas is given as under:

1. **Descriptive Analysis:** Measures of central tendency such as mean standard deviation, and coefficient of variation were worked out to
study the nature and distribution of different variables. These are explained as under:

(a) **Mean:** Mean values have been used to find the average of various items. The following has been used to calculate the arithmetic mean:

$$\bar{X} = \frac{\sum X}{N}$$

Where, $X =$ Arithmetic mean

$\sum X =$ Sum of all the values of the variable $X$, and

$N =$ Number of observations.

(b) **Standard Deviation:** Standard Deviation measures the absolute dispersion or variability from the mean values. The standard deviation values have been calculated by using the following formula:

$$\sigma = \sqrt{\frac{\sum x^2}{N}}$$

Where, $\sigma =$ Standard Deviation

$x = X - \bar{X}$

$N =$ Number of observations.

(c) **Coefficient of Variation (CV):** It is a relative measure of dispersion based on standard deviation. Coefficient of variation was used to test the consistency. There is an inverse relationship between the coefficient of variation and consistency. The series having greater CV are said to be more variable than the other and the series having less CV are said to be less variable than the other.

$$C.V. = \frac{\sigma}{\bar{X}} \times 100$$

Where, $C.V.$ = Coefficient of Variation

$\sigma =$ Standard Deviation, and

$\bar{X}$ = Mean.
(d) **Maximum, Minimum and Range:** Maximum refers to the highest value of an item of a series, whereas minimum is the lowest value of an item of a series. Range is the simplest possible measure of dispersion; and it is defined as the difference between the values of the extreme items of a series.

\[
\text{RANGE} = \text{Highest value of an item of a series} - \text{Lowest value of an item of a series}
\]

(e) **Skewness and Kurtosis:** Skewness measures asymmetry and shows the manner in which the items are clustered around the average. It helps in studying the formation of series and can have the idea about the shape of the curve, whether normal or otherwise. Kurtosis is the measure of flat-toppedness of a curve.

2. **One-Sample Kolmogorov-Smirnov Test:** One-sample K-S Test is used to check the normality of data. The variables having p-value greater than 0.05 are said to be normal.

3. **Exponential Growth Rate (EGR):** Exponential growth rate reflects the strength of movement of any variable over the entire period covered by the study. The exponential function fits in as follows:

\[
Y = ab^x
\]

This function when translated into logarithmic form, gives a log-linearity function.

\[
\log Y = \log a + x \log b
\]

To obtain the value of constants ‘a’ and 'b’, the two normal equations to be solved are:

\[
\sum \log Y = N \log a + \log b \sum X
\]

\[
\sum (X \log Y) = \log a \sum X + \log b \sum x^2
\]

Where, \( a = y\)-intercept

\( b = \) slope of the curve.

When deviations are taken from the middle year then
The above equation takes the following form:

$$\Sigma \log Y = N \log a \text{ and } \Sigma (x \log Y) = \log b \Sigma x^2$$

Therefore, $\log a = \Sigma \log Y/N \text{ and } \log b = (\Sigma x \log Y)/\Sigma x^2$

Antilog of $\log b$ gives the value of $b$. Growth rate is derived from this equation using the following association:

$$b = 1 + r$$

Where, ‘r’ is the exponential growth rate.

4. **Correlation**: Correlation analysis is made to determine the degree of relationship between two or more variables. It does not tell about cause and effect relationship. The values of coefficient of correlation lie between +1 to −1. When $r = +1$, it means there is a perfect positive correlation between the variables. When $r = −1$, it means there is a perfect negative correlation between the variables. When $r = 0$, it means no relationship between the two variables.

$$r = \frac{\Sigma xy}{N \sigma_x \sigma_y}$$

Where, $x = (X-X); \ y = (Y-Y)$

$\sigma_x = \text{Standard deviation of series X}$

$\sigma_y = \text{Standard deviation of series Y}$

$N = \text{Number of pairs of observations}$

$r = \text{the correlation coefficient}$.

5. **Step-Wise Multiple Regression Analysis**: Regression analysis had been applied to study the relationship of independent variables with dependent variable. If there is only one dependent variable and one independent variable used to explain the variation in a dependent variable, then the model is known as simple regression. If multiple independent variables are used to explain the variation in a dependent variable, it is called multiple regression model. The following equation has been used for this purpose:
\[ Y = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n + \varepsilon \]

\( Y = \) Dependent Variable  
\( a = \) constant  
\( x_1, x_2, \ldots, x_n \) are independent variables  
\( b_1, b_2, \ldots, b_n \) are the coefficients of independent variables.  
\( \varepsilon = \) error term

When all the independent variables are not equally important and the correlation among the independent variables is strong then step-wise multiple regression method has been frequently used. The method begins by entering into the model that has the strongest positive or negative correlation with the independent variable and at each subsequent step at the variable with the strongest correlation are entered. In step-wise, at each step the variables are tested for removal.

6. **Analysis of Variance (ANOVA), ONE-WAY AND TWO-WAY**

Analysis of variance (ANOVA) has been carried out to compare more than two means at a time. One-way analysis of variance involves only one categorical variable or a single factor, whereas in two-way analysis of variance, two factors on the dependent variable are studied. The process of analysis is given hereunder:

**One-Factor ANOVA (F-statistics):**

<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>Degrees of Freedom (df)</th>
<th>Sum of Squares</th>
<th>Mean Square (Variance)</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among (Factors)</td>
<td>( c - 1 )</td>
<td>SSA</td>
<td>MSA = SSA/(c – 1)</td>
<td>MSA/MSW</td>
</tr>
<tr>
<td>Within (Error)</td>
<td>( n - c )</td>
<td>SSW</td>
<td>MSW = SSW/(n – c)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( n - 1 )</td>
<td>SST = SSA + SSW</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Where, \( n \) = total number of observations in all groups

\( c \) = the number of groups

\( c - 1 = df_1 \)

\( n - 1 = df_2 \)

MSA is the mean squares among or between variances.

MSW is the mean squares within or error variances.

\[
F = \frac{MSA}{MSW}
\]

Two-way ANOVA (F-statistics):

<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>Degrees of Freedom (df)</th>
<th>Sum of Squares</th>
<th>Mean Square (Variance)</th>
<th>F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor A (Row)</td>
<td>( r - 1 )</td>
<td>SSA</td>
<td>MSA = SSA/(r - 1)</td>
<td>MSA/MSE</td>
</tr>
<tr>
<td>Factor B (Column)</td>
<td>( c - 1 )</td>
<td>SSB</td>
<td>MSB = SSB/(c - 1)</td>
<td>MSB/MSE</td>
</tr>
<tr>
<td>AB (Interaction)</td>
<td>((r - 1)(c - 1))</td>
<td>SSAB</td>
<td>MSAB = SSAB/(r - 1)(c - 1)</td>
<td>MSAB/MS</td>
</tr>
<tr>
<td>Error</td>
<td>( r^c(n' - 1) )</td>
<td>SSE</td>
<td>MSE = SSE/r^c (n' - 1)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( r^c \ast n' - 1 )</td>
<td>SST</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where, \( r \) = the number of levels of factor A

\( c \) = the number of levels of factor B

\( n' \) = the number of values (replications) for each cell

\( n = \) total number of observations in the experiment

SSA means sum of squares due to factor A

SSB means sum of squares due to factor B

SSAB is sum of squares due to interaction of A and B

SSE is sum of squares error

SST is sum of square total
MSA is mean of squares due to factor A
MSB is mean of squares due to factor B
MSAB is mean of squares due to interaction of A and B
MSE is mean of squares error.

Degree of freedom is always add up
rcn'−1 = rc(n'−1)+(c−1)+(r−1)+(c−1)(r−1)

Total = error + column + row + interaction

7. **Reliability Analysis:** Prior to the analysis of results, the research instrument was tested for its reliability and validity. The internal consistency of the grouping of the items was estimated using reliability co-efficient called Cronbach’s alpha and alpha value of 0.70 or above is considered to be the criterion for demonstrating internal consistency of new scales and established scales respectively. The reliability and validity of the responses on the opinion statements have been checked through a scale, viz. reliability analysis. The statements are considered reliable, if the value of Cronbach’s alpha coefficient is 0.70 or higher; and the statements are considered valid, if the mean value of inter-item correlation is 0.20 or higher.

8. **Simple Percentage Analysis:** Simple percentage analysis is the most important and widely used statistical tool in analysis and interpretation of data. The simple percentages are calculated through the following formula:

\[
\text{Single Unit in a Whole of N Units} \times 100
\]

9. **Chi-square Test:** In case of cross tabulation featuring two variables, a test of significance called Chi-square test is used to determine whether the two variables are associated with each other or not. It helps in finding the association between two or more attributes. It has been worked out as follows:
\[ x^2 = \sum \left[ \frac{(O - E)^2}{E} \right] \]

Where, \( O \) = Observed frequencies
\( E \) = Expected frequencies

The calculated value of \( x^2 \) is compared with the table value, for given degree of freedom at a specified level of significance (5% and 10%). If the calculated value of \( x^2 \) is more than the table value, then difference between the variables is considered to be significant or otherwise insignificant.

10. **Average Weighted Score**: Average weighted score has been used to study the borrowers’ perception and views expressed in terms of ranks of preferences for different attributes relating to loan services provided by commercial banks in northern region according to their degree of importance. Five-point Likert Scale has been selected to measure the extent of agreement and level of satisfaction. The range of scale is 1 to 5. The weighted average score has been calculated by assigning weights like 5 to highly satisfied, 4 to satisfied, 3 for neither satisfied nor dissatisfied, 2 to dissatisfied and 1 to highly dissatisfied. On the basis of frequency of ratings for each attribute, average weighted scores have been calculated with the help of following formula:

\[ W = \frac{\sum wfw}{\sum fw} \]

\( W \) = Average weighted score
\( W \) = Weights given to the attribute
\( f \) = Number of respondents who attached weights to the attribute.

11. **Kendall’s Coefficient of Concordance (W)**: Kendall’s coefficient of concordance test has been applied to study the level of concordance among the responses of borrowers about their opinion given by them to various statements. It is considered as an appropriate measure of
studying the degree of association among three or more sets of ranking.

\[ W = 5 \left( \frac{1}{12} \right) K^2 (N^3 - N) - \Sigma T \]

K = Number of sets of ranking, i.e., Number of bank groups.
\[ \Sigma R_j = \text{Sum of ranks assigned by all judges.} \]
\[ \Sigma T = \text{Sum of correction factor for tied observations.} \]
N = Number of objectives to be ranked.
\[ S = \Sigma (R_j - \bar{R}_j)^2 \]

If N is greater than 7, we may use \( X^2 \) value to be worked out as

\[ X^2 = K (N-1) W \text{ with d.f. } = (N-1) \text{ for judging W’s significance at a given level in the usual way of using } X^2 \text{ values.} \]

12. Kruskal Wallis Test (H-Test): It is used for comparing more than two samples that are independent, or not related. The method is used when the data set is skewed. There is a significant difference between the samples when p-value is less than 0.05.

\[ H = \frac{12}{N(N+1)} \left( \frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \ldots + \frac{R_k^2}{n_{kn}} \right) - 3(n + 1) \]

when \( n_1, n_2, \ldots, n_k \) are the numbers in each of k samples.
N = \( n_1 + n_2 + \ldots + n_k \) and \( R_1, R_2, \ldots, R_k \) are the rank sums of each sample.

3.10 LIMITATIONS OF THE STUDY

No doubt, sincere efforts have been made to make the present study a representative one in its related area, but still some limitations have been encountered which are as follows:

1. The secondary data collected for the study carries all the limitations inherent in such data.

2. The present study excluded Regional Rural Banks.
3. Only quantitative factors have been used for the purpose of this study.

4. The primary data in the study has been collected through a pre-designed questionnaire which carries all the limitations inherent with the primary data as perceptions by the knowledge, experience and the attitude of the individuals.

5. As the size of population is very large, therefore, the sample drawn represents the whole population of northern region. So, the shortcomings inherent in this method of sampling may creep into the sample used in the study.

6. As the primary study is restricted to only northern region of India, its findings may not be generalized for other regions of the country as a whole.