CHAPTER- 3

RESEARCH METHODOLOGY

Chapter Outline:

3.1Sample Selection
3.2Data Source
3.3 Study Period
3.4 Tools and Techniques of the Study
3.4.1: Estimation of Trend Growth Rates
3.4.2: Ratio Analysis of Financial Performance Indicators
3.4.3: Fisher's t Test
3.4.4: One -way ANOVA - F test
3.4.5: Factor Analysis
3.4.5.1:Principal Component Analysis (PCA)
3.4.6: Data Envelopment Analysis (DEA)
3.4.7: Cross Tabulation Analysis
3.4.8: Stochastic Frontier Analysis (SFA)

Chapter – 3 Research Methodology

In this chapter the tools and techniques to carry out the study have been described. Apart from this, sample selection, data source, and study period have also been discussed in this chapter.

3.1 Sample Selection:

From the list of NBFCs that are registered with RBI (Assets Finance Companies and Investment Companies) as on 30th November, 2015, a sample of 18 NBFCs on the basis of availability of data for the relevant time period have been selected for our study. Out of these 18 selected companies, 5 companies are investment companies and the rest, i.e., 13 companies are assets finance companies.

The nature of the activities of the companies that are registered as Assets Finance Companies and Investment Companies with RBI are clearly mentioned at the time of registration with RBI. As such, these two sets of companies are mutually exclusive in nature.

3.2 Data Source: The data have been collected mainly from secondary sources, i.e., from the website of RBI and from the published as well as unpublished annual reports of the selected companies. Apart from that, data from some annual reports are collected personally by visiting the offices of the companies.

3.3 Study Period: The period under study spreads over nine years from 2006-07 to 2014-15. The study period starts from 2006-07, since the new category of classification of NBFCs by the RBI was effected in December, 2006.

3.4 Tools and Techniques of the Study:

3.4.1: Estimation of Trend Growth Rates: To calculate the trend growth rates of the selected performance indicators, semi-log trend equation has been used in the study. The semi-log model has been selected since it gives the growth rate directly at a point of time.

Semi- log trend line equation:

$$\text{Log } Y = a + bt + Ut$$

where Y represents dependent variable, a represents constant, b represents growth rate (beta co-efficient), t represents time and Ut represents random disturbance term. In our study Y indicates performance indicators in terms of Share Capital, Reserve & Surplus, Long Term Loan (Liabilities), Short Term Loan (Liabilities), Provisions, Other Liabilities, Fixed Assets, Investments, Long Term Loan (Assets), Short Term Loans (Assets), Cash & Bank Balances, Other Assets, Return on Assets (ROA), Return on Capital Employed (ROCE), Return on Equity (ROE), Debt Equity Ratio (D/E Ratio), Net Profit Ratio (NPR) and Current Ratio (CR).

3.4.2: Ratio Analysis of Financial Performance Indicators: Based on literature review and the nature of the NBFCs, we have identified the following widely used financial ratios and used in our study.

$$\text{ROA} = \frac{\text{Net Profit after Tax}}{\text{Tangible Fixed Assets}}$$

ROA indicates whether or not the fixed assets have been effectively utilized in the operations of NBFCs.

$$\text{ROCE} = \frac{\text{Net Profit after Tax}}{\text{Equity Capital} + \text{ Reserve \& } Surplus + \ Long\ Term\ Debt}$$

ROCE indicates the efficiency of employing long term fund by the stakeholders and the owners of the firm. Thus, it is the basis of the capital employed of the NBFCs which test the profitability related to the sources of long-term funds. The higher the ratio, the more efficient is the use of capital employed. Another important measure of profitability is the return on equity (ROE) defined as

$$\text{ROE} = \frac{\text{Net Profit after Tax}}{\text{Equity Capital} + \text{ Reserve \& } Surplus}$$

The ROE is a measure of the profitability of the NBFCs in relation to the funds supplied by the stakeholders and owners taken together; the return on equity measures exclusively the return on the owner's funds.

Net profit ratio (NPR), another important measure of profitability, is defined as

$$\text{NPR} = \frac{\text{Net Profit after Tax}}{\text{Total Income}}$$

NPR measures the relationships between the net profit after tax and total revenue of the NBFCs. Basically, it indicates the ability of the NBFCs to effectively operate the financing activities during a particular period of time.

Another important ratio that is used as a measure of long term financial performance of the NBFCs is Debt Equity Ratio (DER). It is express as

$$\text{Debt Equity Ratio} = \frac{\text{Long Term Debt}}{\text{Equity Capital} + \text{ Reserve \& } Surplus}$$

It is a measure of long term solvency of the NBFCs. This ratio reflects the relative claims of stakeholders and shareholders against the assets of the NBFCs.

To examine the short term financial liquidity position of the companies, Current Ratio (CR) is widely used. It is expressed as

$$\text{CR} = \frac{\text{Current Assets}}{\text{Current Liabilities}}$$

It is a measure of short term financial liquidity of the NBFCs and indicates the sum of the rupees of current assets available for each rupee of current liability obligations.

Higher the current ratio, the more is the NBFC's ability to meet current obligations and the greater is the safety of funds of short term borrowings.

3.4.3: Fisher's t Test: As in our study, the sample size is less than 30 we have employed the Fisher's t Test to test the difference of mean of selected performance indicators between Investment Companies (aggregative) and selected Assets Finance Companies (aggregative).

The 't' statistic is defined as:

$$t = \frac{\bar{x}_1 - \bar{y}_1}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where n_1 = Sample size of Investment Companies

n_2 = Sample size of Asset Finance Companies

$\bar{x}_1$= Mean of Performance Indicator of Investment Companies

$\bar{y}_1$= Mean of Performance Indicator of Asset Finance Companies

s_1^2 and s_2^2 are the variance of the sample of Investment Companies and Asset Finance Companies respectively.

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

Degree of Freedom = $n_1 + n_2 - 2$

If the observed value is greater than table value (t value) then we reject the null hypothesis and conclude that there is significant difference between two means, while we accept the null hypothesis if the observed value is smaller than the tabulated value.

3.4.4: One-way ANOVA- F test: To test whether the variations in the selected performance indicators for the selected companies under each category of NBFCs are statistically different from each other or not, we have employed one-way ANOVA i.e., F test.

The F-statistic is given by

$$F = \frac{S_1^2}{S_2^2} = \frac{Variance\ Between\ Groups}{Variance\ Within\ Groups}$$

If the observed value of F is greater than the tabled value of F, then we reject the null hypothesis and conclude that there is significant difference in two means, otherwise not.

3.4.5: Factor Analysis: We have carried out factor analysis with respect to selected profitability ratios in order to identify the most important factor or ratio that is the principal factor explaining the variations in the aggregate profitability performance indicator of the selected companies under study.

There are two basic methods of factor analysis, i.e. principal component analysis (PCA) and Common Factor Analysis (CFA). Basically, factor analysis involves techniques which help to create a smaller number of linear combinations on variables so that the contracted variables obtained in this way account for and explain most of the variance in correlation matrix pattern. Principal component analysis (PCA) is an approach to factor analysis that takes into account the total variance in the data, which, unlike the CFA, converts the original variables, 'X_i's into a smaller set of linear combinations 'Z_i's such that the newly formed variables 'Z_i's are independent of one another. In our study, we have employed the principal component analysis for reduction in the volume of data.

3.4.5.1: Principal Component Analysis (PCA): It is a statistical technique that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (entities with various numerical values) into a set of values of linearly uncorrelated variables called principal components. This transformation is described in such a way that the first principal component has the largest possible variance and each following component in turn has the highest variance possible under the constraint that it is orthogonal to the antecedent components.

If we have p variables - $X_1, X_2, \ldots, X_P$, measured on a large sample of n subjects, the i^{th} principal component, Z_i can be written as a linear combination of the original variables 'X_i's. Thus,

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \ldots + a_{ip}X_p$$

So, from the above equation, the linear combination of the first principal component is written as follows

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1p}X_p$$

The above linear combination accounts for the variation in the data (i.e. in the original variables) as much as possible, subject to the constraint that

$$a_{11}^2 + a_{12}^2 + \ldots + a_{1p}^2 = 1$$

Similarly, for the second principal component, the general equation becomes

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \ldots + a_{2p}X_p$$

The second principal component is chosen such that its variance is as high as possible. A similar constraint applies, namely,

$$a_{21}^2 + a_{22}^2 + \ldots + a_{2p}^2 = 1$$

Another constraint is that the second component is uncorrelated with the first component. Rests of the principal components are chosen in the same way.

In PCA analysis, the eigen values are variances of the principal components and the first eigen value is the variance of the first principal component and the second eigen value is the variance of the second principal component, and so on. After the calculation of the principal components it is necessary to decide how many of them will be kept. Obviously, any principal components that account for only a small proportion of the variation in the data (i.e., those with small eigen values) are rejected.

Here, the following points are important to note:

i. The selected principal component is sufficient to account for a particular proportion (e.g., 0.75) of the total variability in the data.
ii. Select only those principal components which have eigen values more than 1.
iii. To form the scree plot of the eigen values which will indicate whether there is an obvious cut-off between large and small eigen values.

3.4.6: Data Envelopment Analysis (DEA): We have applied the method of Data Envelopment Analysis (DEA) to the selected Investment Companies and Asset Finance Companies according to their physical performances.

Data envelopment analysis (DEA) is a non-parametric method used empirically to measure productive efficiency of decision making units (or DMUs).

DEA was designed by Charnes and Cooper to evaluate the relative efficiency of the similar decision-making units. It is a non-parametric optimization method based on mathematical programming. Selection of the efficient units by using this method also guides the inefficient decision-making units with the help of certain reference groups to be efficient. Although, there are many DEA models, the most popular ones are Charnes-Cooper-Rhodes (CCR) and Banker-Charnes-Cooper (BCC) models. CCR models apply constant return to scale assumption, while BCC models use variable return to scale assumption. Each model has two approaches as input-oriented and output-oriented. Input-oriented models analyse the optimal input combination to obtain defined output combination most effectively, while output-oriented ones give emphasis on production of the optimal output combination using defined input combination.

Parameter X_{ij} shows the ith input variable which is used by jth decision-making unit and parameter Y_{rj} shows the rth output variable which is used by jth decision-making unit where $0<I \leq m$ and $0<r \leq s$. v_{ik} and u_{rk} show the weights for input i and output r for the kth decision-making unit respectively. Objective function and constraints of an input-oriented CCR model can be defined as follows:

$$\max h_k = \frac{\sum_{r=1}^{s} u_{rk} Y_{rk}}{\sum_{i=1}^{m} v_{ik} X_{ik}}$$

$$\frac{\sum_{r=1}^{s} u_{rk} Y_{rj}}{\sum_{i=1}^{m} v_{ik} X_{ij}} \leq 1 \; ; \quad j = 1, \dots, n$$

$$u_{rk} \geq 0 \; ; \quad r = 1, \dots, s$$

$$v_{ik} \geq 0 \; ; \quad i = 1, \dots, m$$

where n is the number of decision-making unit.

The fractional model is transformed into input-oriented CCR Primal Model by Charnes and Cooper. Objective function and constraints are given as follows,

$$\max w_k = \sum_{r=1}^{s} u_{rk} Y_{rk}$$

$$\sum_{i=1}^{m} v_{ik} X_{ik} = 1$$

$$\sum_{r=1}^{s} u_{rk} Y_{rk} - \sum_{i=1}^{m} v_{ik} X_{ik} \leq 0 \; ; \quad j = 1, \dots, n$$

$$u_{rk} \geq 0 \; ; \quad r = 1, \dots, s$$

$$v_{ik} \geq 0 \; ; \quad i = 1, \dots, m$$

Variables and DMUs of the DEA Model in our study:

<u>No. of DMUs</u>: In our study, we have carried out the DEA with a total of 18 DMUs (i.e. NBFCs) under two broad categories of DMUs i.e. Investment Companies (5 DMUs) and Asset Finance Companies (13 DMUs).

<u>Output Variables</u>:

In our study, we have taken the following three output performance parameters:

1. Return on Capital Employed (ROCE) which indicates the overall profitability of the DMUs.
2. Debt Equity Ratio (D/E Ratio) which indicates the long term solvency of the DMUs.
3. Current Ratio (CR) which indicates short term solvency of the DMUs.

<u>Input Variables</u>:

In our study, considering the nature of the NBFC regarding the input variables, we have categorized two groups of inputs. One group contains non-revenue items contained in the balance sheet which can be termed as Financial Health Components and other group contains revenue items contained in the Profit and Loss A/C which can be termed as Earning Components. Following are the details of different input variables categorized under the two groups, namely, Financial Health Components (FHC) and Earning Components (EC).

1) Financial Health Components (FHC):

    i) Log of Short Term Loan (Assets),
    ii) Log of Long Term Loan (Assets),
    iii) Log of Short Term Loan (Liabilities),
    iv) Log of Long Term Loan (Liabilities),

v) Log of Value of Investment,

vi) Log of Net Worth,

vii) Log of Cash & Bank Balances, and

viii) Log of Tangible Fixed Assets.

2) Earning Components (EC):

i) Log of Interest Paid,

ii) Log of Total Revenue, and

iii) Log of Employee Cost.

To reduce the number of input variables, we have applied Principal Component analysis (PCA) before applying DEA.

3.4.7: Cross Tabulation Analysis: In our study, we have carried out a cross tabulation analysis to find out relationship, if there is any, between selected Investment Companies (IC) and Asset Finance Companies (AFC) and between their efficiency scores, as derived by DEA during the period under study.

Distinct nominal variables that categorize the characteristics of each observation in a sample of events may be tabulated in a contingency table to show the frequency of co-occurrence of the mutually exclusive characteristics of each variable, as labeled by the rows, columns, and other layers of the cross tabulation. The expected probability of all occurrences in any cell of a cross tabulation of nominal variables (under the assumption of independent probabilities for each variable) is defined as the product of the probabilities of this type of event for each variable. In cross tabulation, the null hypothesis is that the variables tabulated are statistically independent, that is, there is no relation between them. The likelihood of this occurrence is known as a test of statistical significance.

3.4.8: Stochastic Frontier Analysis (SFA): In our study through SFA, we have determined the factors of efficiency and inefficiency as a whole under the two selected categories of NBFCs and the overall efficiency level and the percentage as a whole to increase under the given input and output combinations to make the DMUs efficient. SFA has enabled us to capture the effect of external factors, i.e. stochastic impulse.

Stochastic frontier analysis (SFA) is a method of economic modeling. 'Practically no economic agent can exceed the ideal frontier'; this theoretical idea is the basis of SFA model and deviations from this extreme idea represent individual inefficiencies. The literatures from different perspectives distinguish between production and cost frontiers. The former represents the maximum amount of output that can be obtained from a given level of inputs, while the latter characterizes the minimum expenditure required to produce a bundle of outputs given the prices of the inputs used in its production.

The stochastic frontier model (SFA) was originally developed by Aigner, Lovelland Schmidt (1977). Typically, the production or cost model is a Cobb–Douglas type function, given by

$$\log y = \beta' x + v - u$$

Where y is the observed outcome; $\beta' x + v$ is the optimal production frontier (e.g., maximum production output or minimum cost); $\beta' x$ is the deterministic part of the frontier; and $v \sim N(0, \sigma_v^2)$ is the stochastic part, respectively. The components of x are generally logs of inputs for a production model or logs of output and input prices for a cost model, or their squares and/or cross products. These two parts constitute the stochastic frontier. The amount by which the observed individual fails to reach the optimum (the frontier) is u, known as inefficiency, where $u = |U|$ and $U \sim N[0, \sigma_u^2]$. The SFA model, therefore, can be rewritten as

$$y = \beta' x + v - u \ , u = | U |$$

In the stochastic frontier model, the error term ε is made up of two independent components, *v - u,* where *u* measures technical inefficiency, namely, the shortfall of output *y* from its maximal possible value given by the stochastic frontier $[g(x_0, \beta) + v]$.

When a model of this form is estimated, the obtained residuals i.e., $\hat{\varepsilon} = y - g(x - \hat{\beta})$ may be regarded as estimates of the error term *ε*. The conditional distribution of *u* given *ε,*$E[u|\varepsilon]$ is the mean productive efficiency. Under each of the assumed possible distributional forms for the inefficiency term in a model, this means that this distribution contains whatever information ε yields about u. The predicted value is $\beta' x$. The residual is computed by the Jondrow et al (1982) formula:

$$E[u \mid v - u] \text{ or } E[u \mid v + u]$$

or

$$\hat{E}[u \mid \varepsilon] = \frac{\sigma\lambda}{1 + \sigma^2}\left[\frac{\phi(z)}{1 - \Phi(z)} - z\right], \ \varepsilon = v \pm u, \ z = \frac{\varepsilon\lambda}{\sigma}$$

The marginal effects in the model are the coefficients β. Estimation and analysis of the inefficiency of individuals in the sample and of the aggregated sample have greater influence on the model than evaluation of the model parameter. The results obtained in this way are critically dependent on the model form and the assumptions set. In order to overcome this, special focus has been given to panel data estimation technique.

3.5 References:

1. Aigner, D. J.; Lovell C. A. K.; and Schmidt, P. (1977), "Formulation and estimation of stochastic frontier production functions", *Journal of Econometrics,* pp. 6:21-6.37.
2. Banerjee, Bhabatosh (2008), Fundamentals of Financial Management, PHI Learning Private Limited, New Delhi, 1st Edition, pp. 431-466.
3. Banker, R. D., Conrad R. F., and Strauss, R. P. (1986), "A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production", *Management Science* 32(1), pp. 79-96.
4. Bhattacharyya, D. K. (2006), Research Methodology, Excel Books, New Delhi, 2nd Edition, pp. 296-311.
5. Charnes, A., Cooper, W. W., Seiford, L., and Stutz, J., (1983), "Invariant multiplicative efficiency and piece-wise Cobb-Douglas envelopments", *Operations Research Letters*, pp: 46-62.
6. Charnes, A., Cooper, W.W., and Rhodes, E. (1979), "Short Communication: Measuring Efficiency of Decision Making Units", *European Journal of Operations Research*, pp: 247-251.
7. Damodaran, Aswath (2007), Corporate Finance, Wiley India (P) Ltd., New Delhi, 2nd Edition, pp. 89-107.
8. Das, N. G. (2008), Statistical Methods Vol. II, Tata McGraw Hill Publishing Company Limited, New Delhi, 1st Reprint, pp. 233-237 and 273 -281.
9. Gupta, S. K. and Sharma, R. K. (2011), Management Accounting, Kalyani Publishers, Ludhiana, 12th Edition, pp. 4.1-4.80.

10. Jondrow, J., C. A. Knox Lovell, Ivan S. Materov, and Peter Schmidt (1982), 'On the estimation of technical inefficiency in the stochastic frontier production function model.' *Journal of Econometrics,* pp. 233–238.
11. Karadag, E. and Huseyin, T. (2017), "Data Reduction in Data Envelopment Analysis: A Research on Efficiency of Insurance Companies", Journal *of Current Researches on Business and Economics,* Vol. 7, Issue 1, pp: 66-77.
12. Khan, M. Y. and Jain, P. K. (2015), Management Accounting, Tata McGraw Hill Education (India) Pvt. Ltd., New Delhi, 4th Edition, pp. 6.1-6.88.
13. Kuchler, Andreas (2013), "The efficiency of Danish banks before and during the crisis: A comparison of DEA and SFA", *Danmarks National Bank - Working Papers*, Vol. 87, pp: 4-29.
14. Necmi K. Avkiran (2001), "Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis", *Socio-Economic Planning Sciences,* Vol. 35, pp: 57-80.
15. Panneerselvam, R. (2016), Research Methodology, PHI Learning Private Limited, New Delhi, 2nd Edition, pp. 224-298.
16. White, D. R. (2004), "A Student's Guide to Statistics for Analysis of Cross Tabulations", *World Cultures*, Vol. 14 (2), pp: 179-193.