METHODOLOGY

CHAPTER III

METHODOLOGY

The methodology used for the study is described in this chapter. It includes the focus of the study, selection of banks, sources of data, choice of time series, tools of analysis and the choice of test statistics appropriate for the purpose of analysis.

Focus

The main focus of the study is on performance of the commercial banks in India. Their performance affects on the supply side, the economic growth as measured by real GNP / GDP and on the demand side by the liquidity preference and investment climate. Both demand and supply are conditioned by the expectations of rate of economic growth on one hand and concern for stability on the other, the latter emerging as the target variable in monetary policy. Therefore, Indian economy is the universe for the study, while the commercial banks draw the sectoral focus.

Banks Studied

There are several categories of banks in India. Among them State Bank of India (SBI) is a special category. There are 19 nationalized scheduled banks, 30 other scheduled banks and 33 foreign banks. They are covered by this study in the aggregate and no sample or individual units among them are studied, largely due to the limitation of time. There are 196 Regional Rural Banks (RRB), several non-scheduled banks and cooperative banks. As they function more for specific social purposes rather than on purely commercial purposes, they are excluded form this study, again for want of time and sufficient data to separate their banking and non-banking activities. Thus, the study is based on the aggregate data for the four categories of commercial banks viz., (1) State Bank of India and its associates (SBI), (2) Nationlised Scheduled banks (NSB), (3) Other scheduled banks (OSB), and (4) Foreign banks (FB).

Data

As the study involved mostly macroeconomic variables for the nation as a whole, it is based on secondary data on bank rate, related variables and variables measuring performance of the banks. More specifically the secondary data on the following macro economic variables were collected for the objectives one and two.

1. Gross Domestic Product at factor cost at constant prices *(g)*,
2. Money Stock - Short money (M_1),
3. Money Stock - Broad Money (M_3),
4. Bank rate *(r)*,
5. Cash Reserve Ratio *(CRR)*,
6. Statutory Liquidity Ratio *(SLR)*,
7. Foreign Exchange Reserve *(FOREX)*,
8. Scheduled Commercial Banks *(SCB)*, and
9. Wholesale Price Index (1993-94 =100) for all Commodities *(WPI)*.

For studying the performance of the commercial banks the data on the following variables were collected separately for the four categories of bank viz., SBI, NSB, OSB, FB and for their total. Variables for the objective three are:

1. Number of banks *(No.)*,
2. Aggregate deposits *(AD)*,
3. Liabilities *(L)*,
4. Assets *(A)*,
5. Investments *(I)*,
6. Loans and advances *(L&A)*,
7. Non Performing Assets (NPA),
8. Net profit (π),
9. Credit to assets ratio (CAR), and
10. Priority sector credit to total credit ratio (PSR).

Sources of Data

As the study involved analysis of only macro data for the country as a whole, it is based on secondary data collected mainly from the publications of RBI that include **Statistical Tables Relating to Banks in India, Report of Trends and Progress of**

Banking in India, RBI Bulletins, Report on Currency and Finance, Hand Book of Statistics on Indian Economy, Annual Reports of RBI, and circulars on credit policy. Several other publications such as the **Annual Economic Surveys, CMIE Reports, and Statistical Abstracts of India** were also consulted. Besides serving as sources for relevant data, these publications also provided information about the monetary policy measures undertaken by the RBI from time to time as also the limits prescribed for the expansion of money supply / credit.

Period of Study

The Government of India (GOI) nationalized 14 commercial banks on 19th July 1969. As the experiment was new and commercial banks had to reorient their activities the arrangement got established by 1975. Six more banks nationalized in 15th April 1980. Today there are 19 banks in the category of Nationalized banks. The secondary data for the study were therefore collected for the period from 1975 to 2004.[1] This period marks the post – bank nationalization period. It is further subdivided into two sub-periods.

Sub-Period-I – 1975 to 1990

Sub-Period-II – 1990 to 2004

The GOI implemented a policy of liberalization, privatization and globalization of the economy since mid-1990 aiming at stabilization and structural adjustments in the economy to achieve sustainable and rapid economic growth and it is called new economic reform policy (NERP). Therefore the sub-periods: I and II represented pre- and post- (NERP) reform periods respectively. The two periods are compared for the differences in the key variables studied. Therefore time series of annual data on Bank Rate, money supply (M_3), liquidity, GDP at factor cost and wholesale price index (WPI) were collected for the study period viz., 1975 –2004.

Analysis

Collected secondary data were first processed to make them comparable. Particularly the time series data for some of the indexed variables such as GDP at constant prices, and

[1] The data referred to the accounting period of 12 months, for the fiscal years (i.e) January – December for years 1975 to 1987 and April – March next for the years from 1988-89 to 2003-2004.

whole sale price index were available with 1970-71 as the base for the period up to 1993-94. First attempt was therefore to bring the data to a common base of 1993-94 = 100.

Tools of Analysis

Above variables were first studied for annual average compound growth rate (CGR) separately for pre- and post reform period by using the following function:

$$Y_t = A(1+g)^t \, e^{ut} \qquad \ldots\ldots (1)$$

Taking log on both sides of equation (3.1), we have:

$$\text{Log } Y_t = \log A + t \log (1+g) + u_t \qquad \ldots\ldots\ldots (2)$$

Where,

Y_t = the value of the variable under consideration in period t,

A = a parameter,

g = a parameter that is the compounded rate of growth of Y,

u_t = the disturbance term.

The equation (3.2) can be rewritten as

$$Y_t{}^* = a^* + b^* t + u_t \qquad \ldots\ldots (3)$$

Where,

$Y_t{}^* = \text{Log } Y_t$.

$a^* = \log A$

$b^* = \log (1+g)$

Equation (3.3) is estimated using ordinary least squares method. Since log (1+g) = b*, the estimate of growth 'g' (in percentage) is given as:

$$g = [\text{antilog } (b^*) - 1] \times 100 \qquad \ldots\ldots\ldots (4)$$

Trend Analysis

Next attempt was to study the trend in the reference variables. A linear trend equation was specified. In the trend equation, (binary) dummy variable D was introduced to know the difference between pre-reform and post-reform periods in the trend as a slope dummy variable. D takes value (0,1) for pre- and post-reform periods respectively.

The coefficient of D was expected to have a positive sign. The estimation of trend equation was by ordinary least squares method. With this the specified linear trend equation was:

$$Y_t = a + bt + cDt + e_t \qquad \text{.................(5)}$$

Where,

Y_t = value of the variable studied in the year t,

t = trend variable for t = 1,2, ... 29,

a, b, c = are parameters to be estimated

e_t = random error term.

Test for Auto-regressive Errors

Any time series would require a test for the condition of non – auto regression or absence of serial correlation, because "the assumption of non auto regression was frequently violated in the case of relations estimated from time series data (Kmenta, 1971, p.269)[2]. When this problem is present, the OLS estimators are not efficient estimators and the precision of the test is lost. Therefore is the need to test for the presence are absence of auto regression in the time series data estimation. Conventionally Durbin – Watson (d.w) test is widely used. To apply this test, calculate the value of a statistics 'd' given by,

$$d = \frac{\sum_{t-2}^{n} (e_t - e_{t1})^2}{\sum_{t=1}^{n} e_t^2}$$

Where e_t is the ordinary least square residuals. We test the null hypothesis : $H_o : \rho = 0$ against a one sided or two sided alternative.

If the alternative hypothesis was positive auto – regression, ($H_A : \rho = 0$) the decision rules are:

1. Reject H_o if $d < d_L$
2. Do not reject if $d > d_U$
3. Test is inconclusive if $d_L \leq d \leq d_U$

---

[2] Jan Kmenta (1971), ***Elements of Econometrics***, (New York : Mac Millan Publishing Company), pp.269-275.

If the alternative hypothesis was a two sided one, i.e., if H_A : ρ unrestricted, then the decision rules are;

1. Reject H_o if $d < d_L$ or if $d > 4 - d_L$
2. Do not reject H_o, if $d > d_U < d < 4 - d_U$
3. Test is inconclusive if $d_L \leq d \leq d_U$ or $4 - d_U \leq 4 - d_L$

The values of d_L (for lower limit) and d_U (for upper limit) are given in Durbin – Watson. These values vary with the number of observations (n) and the number of explanatory variables (K) in the regression. The test is not applicable in lagged values of the dependent variable enter as explanatory variables. This case is not involved in this study. For the number of observations (n) = 29 and K up to 5, the d_L and d_U values are presented below.

Table: 1.- Durbin – Watson Test – Limit Values

Significance	**K=1**		**K=2**		**K=3**		**K=4**		**K=4**	
	d_L	**d_U**	**d_L**	**d_U**	**d_L**	**d_U**	**d_L**	**d_U**	**d_L**	**d_U**
5%	1.34	1.48	<u>1.27</u>	<u>1.56</u>	1.20	1.65	1.12	1.74	1.05	1.84
2.5%	1.24	1.38	1.17	1.45	1.10	1.54	1.03	1.63	0.96	1.73
1%	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61

***Note*:** N=29
Source*:** Jan Kmenta (1971), ***Elements of Econometrics, (New York :Mac Millan Publishing Company New York), pp.625-627.

Above test was used for one tailed test at 5% level for n = 29, and K = 2 (values underlined) in this study.

Stationarity of Time Series

The ordinary least squares method of estimation is based on the assumption that the means and variances of the variables are well defined, fixed constants, and are independent of time[3]. However, the econometricians have pointed out that means and variances of most of the macro-economic time series variables change over time and hence are non stationary.[4] In such cases, all the computed statistics in a regression, which

[3] B.B.Rao (1994), ***Cointegration for the Applied Economists***, (New York : St. Martin's Press), p.2.

[4] C.R.Nelson and C.I Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series", ***Journal of Monetary Economics***, 10 : pp.139-162.

use these means and variances, also become time dependent and their values fail to converge as the sample size increases. Furthermore, conventional tests of hypothesis will be seriously biased towards rejecting the null hypothesis of no relationship between the dependant and the independent variable.[5] Thus, the application of classical estimation methods such as ordinary least squares (OLS) gives misleading inferences. For instance, one may obtain a very high value of R^2 while there may not be any meaningful relationship between the two variables.[6] This problem is commonly known as the spurious regression problem. It arises when both the time series exhibit strong trends. The high R^2 observed is due to the presence of trend, and not due to a relationship between the variables.[7]

However, when two non-stationary time series variables combine to generate a stationary error process in a regression, they are said to be co-integrated. If two non-stationary variables are co-integrated, then the inference of a significant link between them might not be spurious at all[8].

Therefore, the first step in the present analysis is the test for unit roots. For this, the Dickey Fuller (DF), and Phillips Perron (PP) tests are used. The second step is to investigate the presence of co-integration between pairs of variables. Dickey Fuller (DF) and Phillips Perron (PP) tests for co-integration were applied for testing the presence of co-integration. The step involves the estimation of ECM.

Dickey Fuller (DF) Test

An appropriate and simple method for testing the order of integration of a stochastic time series was proposed by Dickey and Fuller (1979)[9]. This test examines the presence of a unit root in the following AR (1) model:

---

[5] C.R. Nelson and C.I. Plosser, op.cit., p.2

[6] D.N.Gujrati (1995), ***Basic Econometrics***, (Singapore : MgGraw Hill), p.709.

[7] G.S.Maddala and Kim (1998), ***Unit Root, Cointegration and Structural Change,*** (Macmillan Publishing Company, Singapore), p.28.

[8] Cromwell, et.al. (1994), "Multivariate Tests for Time Series Models", ***Sage University Paper series*** on Quantitative Applications in Social Sciences, 07-100, Thousands Oaks, Canada, pp-17-18.

[9] D.A.Dickey and W.A.Fuller (1979), " Distribution of the Estimators in Autoregressive time Series with a Unit Root", ***Econometrica***, 49 : pp.1057-1571.,

$$Y_t = \rho y_{t-1} + e_t \; t = 1.2......n \qquad ......(6)$$

Where e_t is a set of independent and identically distributed random variables. The null hypothesis is $|\rho| = 1$ against the alternative hypothesis that $|\rho| < 1$.

The above equation is rewritten as:

$$\Delta Y = \alpha_1 Y_{t-1} + \epsilon_t \qquad .... (7)$$

Where, $\rho = 1+ \alpha_1$

Hence to test $|\rho| =1$ is same as to test for $\alpha_1 = 0$. The test involves examining the t-statistic for the parameter α_1 in equation 3.5 under the null hypothesis of Y_t being non-stationary, which implies that it is a random walk or possesses a unit root.

The null hypothesis is rejected if the t-statistic is smaller than the critical value at a specified level of significance. The critical values are obtained by Dicky (1976)[10] using Monte Carlo methods.

The DF test can also be used for testing unit roots for a variable generated as a stochastic process with drift, i.e. by testing the following equations:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t \qquad .... (8)$$

Where, α_0 is a constant representing drift. This means that the mean of Y_t is non-zero. A linear trend (T) can also be added to the DF equation above. This would be a combination of both trend and difference stationary components. The test equation in this case is:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_t + \epsilon_t \qquad .... (9)$$

The t-statistics associated with Y_{t-1} in equations (8) and (9) are called DF-statistics 'without trend' and 'with trend', respectively. It is to be noted that the critical values for the above two cases differ (see Table 2).

[10] D.A.Dickey (1976), Estimation and Hypothesis Testing in Non-Stationary time Series, ***Ph.D Thesis***, Lowa State University, Ames, IA).

Table: 2.- Critical Value for Tests for Unit Roots

Level of significance	1%	5%	10%
Without Trend	-3.43	-2.86	-2.57
With Trend	-3.96	-3.41	-3.13

Note: The Critical values are same for DF, ADF and PP tests.
Source: Computer Package SHAZAM.

Test procedure

Step 1: First equations (8) and (9) were estimated. The null hypothesis for each case is:

(1) H_0 : Y_t is a random walk plus drift, $\alpha_1 = 0$.

(2) H_0: Y_t is a random walk plus drift around a stochastic trend, $\alpha_1 = 0$.

The alternative hypothesis is that $\alpha_1 \neq 0$ in 1 and 2 above.

Step 2: For each of the case in step 1, the t-statistic was calculated as:

$t(\alpha_1) = \alpha_1 / SE(\alpha_1)$

Step 3: The significance level (5 percent) and critical values (under lined) for each case were accordingly chosen.

Step 4: In case, the calculated value of 't' was found to be less than the negative critical value; the null hypothesis of unit root was rejected. Otherwise, the null hypothesis could not be rejected.

Phillips Perron (PP) Test

Phillips and Perron (1988)[11] have suggested a test to overcome the problem of serial correlation in the DF test equations. They have proposed estimating the original DF equation by OLS even when e_t is serially correlated and then modifying the various statistics to take account of serial correlation (Hamilton (1994)[12] and Maddala (1998)[13])

---

[11] P.C.B. Phillips and P.Perron (1988), "Testing for Unit Root in Time Series Regressions", ***Biometrica,*** 75, p.335-346.

[12] J.D. Hamilton (1994), ***Time Series Analysis***, (New Jersey : Princeton University Press, Princeton), pp.506-514.

[13] G.S. Maddala and Kim (1998), ***Unit Root, Cointegration and Structural Change,***(Singapore : MacMillan Publishing Company), pp.78-81.

Unit Root Tests (DF and PP) With Trend (Level Data in Logarithm)

As noted by Holdon and Perman (1994),[14] the unit root tests without trend cannot distinguish between trend stationary and difference stationary processes. Thus, the DF and PP tests were applied by including a linear trend variable also in the corresponding test equations.

In order to examine whether the variables possess one or more unit roots, the above tests were applied to the first difference of logarithm of the data (hereafter called as the first differenced data) of the variables.

Performance of the Commercial Banks

It was studied for key variables: net profit, profitability, deposits, advances, investments, total assets, net worth, spread (No. of branches) and growth rate (CGR) by using function as above (3.1 to 3.4) and non–performing assets. Simple tabular analysis was used for a comparative study of the four categories of banks for the selected years: 1975, 1990, and 2004 (latest period for which data were available). Several financial ratios also were studied for the above years, and they were / ratios of:

1. Net Profit to Total Assets
2. Total advances plus Investments to total Deposits
3. Net NPA to Total Assets
4. Net NPA to Net Advances
5. Capital Adequacy Ratio
6. Cost to Income Ratio

Profit Function

Profit or profitability of banks is a summary measure of their performance, because profit in banks acts as a stimulant to the management to expand and to improve their services. Adequacy of profit is important for the viability of the banks because, increasing pre–emptions for cash reserve ratio (CRR), statutory liquidity ratio (SLR), the rigorously structured interest rates, enormous increase in the establishment cost and

[14] D.Holden and R.Permen (1994), "Unit Roots and Cointegration for the Economists", in B Bhaskara Rao (ed), ***Cointegration for the Applied Economist,*** (New York : St.Martin's Press), pp.47-95.

the burden of social banking have serious effect on the profit / profitability of the banks in India. Therefore an attempt was made to identify the major determinants of profit / profitability of the banks in India, with the help of a profit / profitability function. A risk and equity adjusted Profit Function was specified on the basis of Waraptr[15] and Elavia Bansal's[16] models:

Profit Function Model

$$\begin{aligned} \ln \pi_t = {} & \alpha_o + \alpha_1 \ln c_{it} + \alpha_2 \ln c_{2t} + \beta_1 \ln P_{it} + \beta_2 \ln P_{2t} \\ & + r_1 \ln N_t + r_2 \ln FAPB_t + r_3 \ln TDPB_t \\ & + \delta_1 \ln R_{it} + \delta_2 \ln R_{2t} + \delta_{3t} \ln R_{3t} + \delta_4 \ln R_{4t} \\ & + \emptyset_1 \ln SCR_t + \emptyset_2 \ln MS_t + \emptyset_3 D + \mu_t \end{aligned}$$

Where,

ln = shows that the values are in log to base e

π_t = net profit (in Rs. crores)

C_{it} = input price on B_1 (i.e) b_1 / B_1(in Rs.)

β_{it} = total deposit + borrowing – current deposits (in Rs.crores)

b_{it} = interest paid on B_1 (in Rs.)

C_{2t} = input price on B_2 (i.e) b_2 / B_2 (in Rs.'000/ worker)

β_{2t} = number of employees

b_{2t} = wages and salaries paid for B_2 (in Rs. crores)

P_{it} = output price of A_1 (i.e) a_1 / A_1, (in Rs), Where

A_{it} = total loans and advances, investments and bills purchased and discounted (in Rs. crores)

a_{it} = interest earned from A_1 (in Rs. crores)

P_{2t} = output price of A_2 (i.e) a_2 / A_2, (in Rs. crores), Where

A_{2t} = other earning Assets / Services (in Rs.crores)

a_{2t} = income from A_2 (in Rs.crores)

---

15 I.Waraptr (1983), "The Risk Adjusted Profit Funtion: Measurement of Economies of Scale and Efficiency in Commercial Banks", ***Ph.D Thesis***, University of Illinois, USA.

16 B.H.Elivia and N.Sharad Bansal (1993), "Economies of Scale in the Indian Banking Industry: A profit Function Approach", ***Journal of Indian Institute of Bankers***, 64 (1):22-36.

N_t = number of branches

$FAPB_t$ = fixed assets per branch (in Rs. crores)

$TDPB_t$ = total deposits per branch (in Rs. crores)

R_{it} = ratio of total credit to total assets

R_{2t} = ratio of total borrowings to total liabilities

R_{3t} = burden@ as percentage to total assets (%)

R_{4t} = ratio of priority sector credit to total bank credit

SCR_t = SLR+CRR (%)*

MS_t = market share (%) of deposits of particular group to total bank deposit in the market (excludes deposits with RBI) for group *j* for j = 1,2,3,4. This variable drops out in the function for aggregate of all groups.

D = a dummy variable with value zero for pre-reform period (1975-1990) and 1 for the post-reform period (1991-2004). It is an intercept dummy.

μ_t = Stochastic error term

t = 1,2……n years for which data were used. It is time series of data for the years from 1975 to 2003-04.

@ - Burden is defined as the difference between non-interest expenditure and non-interest income.

* - The incremental pre-emptions is simply sum of CRR, incremental CRR and SLR.

Parameters to be estimated and their sign expected *a priori*

$\alpha_o > 0, \alpha_1, \alpha 2 < 0 ; \beta_1, \beta_2 > 0 ; r_1, r_2, r_3 > 0$

$\delta_1, \delta_2, \delta_3$ *unrestricted*, $\delta_4 < 0$

$\O 1 < 0, \O_2 < 0, \O_3 > 0$

In the model, when π was replaced by Y where Y was net profit as a percentage to total assets, then it became profitability function, with the same set of explanatory variables. Both profit (π) and profitability (Y) functions were attempted, but the latter had to be dropped due to the problem multicollinearity. The profit function was estimated separately for the four groups of banks *viz.*, $j=1$, State bank and its associates (SBI), j=2, Nationalised scheduled banks (NB), j=3, Other scheduled banks (OSB) and j=4, Foreign banks (FB) and also for the aggregate of all banks treated as group j=5. This allowed a comparative study of the four groups of banks. In all, there were five estimated equations. They were evaluated for drawing inferences with the help of R^2 - the coefficient of multiple determination, t- statistics for the partial regression coefficients and Durbin-Watson (d.w) statistics for the absence of auto regressions and F- statistics. All statistical tests were performed for five percent level of significance or 95 % confidence interval denoted by a single $*$.

The Influence of the Bank Rate

The main focus of this study is on the influence of Bank rate on the performance of commercial banks. However the influence of bank rate is through four performance variables. They are: spread (Number of branches), profit, liquidity, and net worth. With each variable as a dependent variable a multiple regression model was specified. Independent variables were: (1). Bank rate (X_1), CRR (X_2), SLR (X_3), rate of inflation (X_4), GDP (X_5) and Money supply (X_6), Foreign exchange reserve (X_7) and D, as defined above. This multiple regression is necessary because the banks rate does not act in isolation. The time series data for the period from 1975 to 2004 were used to fit the regression equations. Specific attention was on the coefficient ($\beta 1$) of Bank Rate (X_1). The model was:

$$Y_{jt} = \beta_o + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + \beta_6 X_{6t} + \beta_7 X_7 + \beta_8 D_t + e_t$$

for $j = 1, 2, 3, 4$ dependent variables spread, profit, liquidity, net worth

for $t = 1, \ldots 29$ years t_1, being 1975 t_{29} being 2004.

Where,

Y_j = Dependent variable for j = 1, 2, 3, 4

X_1= Bank Rate (%)

X_2 = CRR (%)

X_3 = SLR (%)

X_4 = Inflation (%)

X_5 = GDP (index with 1993-94 = 100) (at factor cost)

X_6 = Money supply index (with 1993-1994 =100)

X_7 = Foreign exchange reserve (Rs. crores)

D_t = 0 for t = 1, ... 15

1 for t = 16, ... 29

e_t = random error term.

After studying for the possible problems of auto-regression and multicollinearity with the help of Zero order correlation matrix, the model was estimated by the ordinary least squares (OLS) method.

Finally the relationship between bank rate *(r)* and other macro variables (g, M_1, M_3, (M_1/ M_3), F_x, WPI, and FD) was studied with the help of simple correlation coefficients.

All statistical tests were done for 5% of significance with the help of appropriate (t, F, R^2, χ^2) test statistics.