CHAPTER VI

DETERMINANTS OF THE STATE LEVEL COMMERCIAL BANK'S SSI CREDIT

6.1 Problem of Standardising and Defining Indicators:

Banks' SSI credit in absolute quantum has varied widely from state to state. However, this is not unexpected as the states themselves are of various types with respect to population size, geographical area and its elongation, volume of economic activities in general and activities relating to SSI in particular, infrastructural facilities, availability of raw materials and, of course, supply potentiality of skilled labours requirement etc. Our problem, therefore, has not been one of explaining the variation in absolute quanta of SSI credit of banks to different states, but of explaining the variation in such credit of commercial banks as a standardised average per unit of requirement of SSI finance expressed by some well-defined indicators or represented by some parameters. The phrase, "requirement of SSI financé, though used freely, requires proper elucidation. It is very difficult to define this term without objective assessment of financial requirements in SI in each state measured in terms of an optimum norm of potentiality and possibility of industries in a small scale.However,such an optimum efficiency norm is yet to be established, at least in India with so much diversity regarding heterogeneity in economic development, culture, historical background, tradition, perception of social objectivity, language, level of education etc. This problem is, of course, not there only for Small Scale Industries in India, but for the objective assessment of requirements of any developmental work.

We, therefore, conceive this efficiency norm as a function of some parameters envisaged from the knowledge of economic principle in general. This facilitates to visualise small scale industry as a dependent variable functionally related with the parameters relating to the following four major non-banking areas:

1. Economic and productive activities in general
2. Infrastructural facility
3. Geo-spatial and natural condition
4. Social and human resources

Economic and productive activity encompasses :

i. Level of industrial activities, such as, number of industrial workers, volume of capital invested, volume of industrial production etc.

ii. Level of agricultural activities, such as, volume of agricultural product, workers involved in agricultural activities, facility of irrigation, extent of mechanisation in agricultural and other factors which influence the intensity of agriculture.

iii. Volume of trade

SSI has bearing on the following infrastructural facility

i. availability of energy like electricity

ii. existence of marketing and distribution channel like rail road, motorable road, air link and provision of frequency of take off, port facility.

iii. Information network etc.

One of the most important factors, social and human resources should enlist,

i. spread of modern institutional education, both in rural and urban area,

ii. literacy rate and educated unemployment, both technical and non-technical,

iii. other civic and urban facilities like health-care, housing facilities etc. in urban and rural areas.

Geo-spatial and natural condition means

i. the proximity of natural resources, like, mining activities, forest resources etc,

ii. scope of power generation hydel or other etc.

Here, also, one is faced with a problem of choice and a data constraint. We have both types of problem, time-series data are not available at state level for some of the factors mentioned above, and remaining factors are either imponderable or can not be quantified. We are, therefore, left with the choice of taking some variables each proximate to a subset of above factors. We have introduced and discuss some banking parameters TC, TD, CDR and BOPTP in the Chapter IV . Along with those we here define two more variables, namely, Urbanisation Index (U_I) and Literacy Index (L_I). Urbanisation Index can be regarded as a proxy variable of factors under first two major non-banking areas and is defined as,

$$U_I = \frac{\text{Percentage of urban population in a state}}{\text{Percentage of urban population in India}}$$

It is perceptible that the the intensity of the first two major factors in a state, more or less, determines the degree of urbanisation. In the absence of appropriate data, this urbanisation index can be a variable well embracing above economic and productive as well as infrastructural factors.

Factors of social and human resources that we have mentioned earlier, would be represented by Literacy Index (L_I) as those having direct consequence on literacy rate. Moreover, banking activity itself depends, to an extent, on the spread of education, as banking - a modern, civilised instrument of economic system, requires a minimum level of literacy as an essential precondition of smooth functioning. In this respect, literacy index would be a close proximation of factors under broad category of human and social situation and is defined as,

$$L_I = \frac{\text{Percentage of literacy in a state}}{\text{Percentage of literacy in India}}$$

Third group of factors i.e., geo-spatial and natural condition, to measure the proximity of natural resources could not be included in our analysis due to non-quantifiablity and subjectivity. Factors, under this area, which are non-quantifiable, either are imponderable in nature or un available. Therefore, we have no choice but to keep it beyond our analysis.

Finally, therefore, we have the following dependent and explanatory variables -

x_1 - Statewise per capita outstanding SSI credit of banks,

x_2 - Statewise per capita outstanding credit of banks,

x_3 - Statewise per capita deposit of banks,

x_4 - bank office per thousand population in a state,

x_5 - Credit-deposit ratio of each state,

x_6 - Literacy Index of a state,

x_7 - Urbanisation Index of a state.

6.2 Regression analysis of per capita SSI unit:

In this section we shall take up the analysis of regression of x_1, step-wise, to estimate to what extent the addition of variables mentioned above, one by one, improve the overall system of determination of the state level of commercial banks' credit to SSI in all-state sample.

We, of-course, should take note of the fact that,in actual financial world, commercial banks' SSi credit is not precisely determined by the six variables taken into account for analysis here. A number of other factors, such as, rate of interest, cost of management of a SSI unit (which determines the economic viability of the project), productivity and cost of production excluding overhead expenditure, rate of repayment and many more intrinsic and ektrinsic factors will cause actual SSI credit given by commercial banks' to a state to deviate from predicted values. However, lack of data did not allow us such a comprehensive exercise. Our contention in this chapter, of-course, also not to list out exhaustive factors influencing SSI credit. Our objective is only to proximate causes influencing SSI credit distribution can be understood in a better way. Hence our function would incorporate above mentioned variables only.

We tried to estimate here the following regression equation for our analysis:

$$x_1 = \lambda_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4 + \lambda_5 x_5 + \lambda_6 x_6 + \lambda_7 x_7 + u \qquad (1)$$

6.2.1 The problem of multi-collinearity & justification of methodology applied:-

An important condition for the application of least squares is that the explanatory variables must not be perfectly linearly correlated. The presence of linear correlation among some or all explanatory variables is expressed by the term "Multi-collinearity". The problem is, in practice, perfect correlation correlation among explanatory variables is not often found. But, on the other hand, in most cases there is some degree of inter-correlation among the explanatory variables. Thus, though the simple correlation between two explanatory variables may not be exactly one even for any value ranging between zero and unity, the Presence of 'multi-collinearity' is possible and it may hinder the accurate and stable estimation of the parameters. At the same time, it is true that the exact effects of collinearity is yet to be established theoretically by the econometricians[1].

Second problem we want to mention is that the presence of multicollinearity would make the estimation of coefficients indeterminate and the standard error, as often said, of these estimates would become infinitely large.

For the both the cases, research findings of applied econometrics are controversial and by no means conclusive. In the first case, if the explanatory variables are not perfectly collinear but are to a certain degree ($0 < r_{ij} < 1$), the effects of collinearity are uncertain[2]. Similarly, for the second case, in some studies the standard errors of the estimates are substantially increased when collinear variables are being introduced in the function, while in other cases, standard errors have not affected by the incidence of multi-collinearity[3]. Multicollinearity, however, is not a problem, as shown by Klein[4], unless the simple correlation between any two explanatory variables is higher than the overall degree of multiple correlation among all variables simultaneously, i.e.,

$$r_{ij}^2 \geq R_{Y.x_1,x_2,\ldots\ldots\ldots,x_k}$$

where r_{ij}^2 is the simple correlation between x_i and x_j. Farrar & Glauber[5], of course, are critical about Klein's view. Theil's[6] proposition, on the contrary, is that even small inter correlation between variables, explanatory variables being more than two, may be the cause of non-significance of estimations due to increase of standard errors. Thus, the large standard error has an impact over estimation and presence of multi-collinearity will not always necessarily show the large standard error, as shown by Frisch[7].

It is generally accepted now that increasing standard errors should be observed, with some exceptions, if we include inter correlated variables as explanatory in the function. Thus, endeavor of estimation of a function with multi-collinearity, has the degree of making mis-specification. A variable, which may be an important determinant of the variation of the dependent variable, may be rejected due to high standard error.

We have calculated partial correlation of dependent and explanatory variables of the function (1) and partial correlation matrices are given in Table 6.1 to 6.17 for each year separately. Simple correlations of x_1 with x_2, x_3, x_4 and x_7 separately showed high degree of correlation. Similarly, correlations are high between x_2 and x_3, x_3 and x_4, x_3 and x_7, x_4 and x_7. But, standard errors of λ_3 did not appear large and, t-values were not significant at their respective degrees of freedom, as shown in the Tables. Whereas, standard errors of λ_4, λ_5 and λ_5 appeared very large as well as t-values were insignificant at 1 percent level.

A close look at the Tables showing partial correlations and standard errors of respective regression co-efficients indicate the presence of multi-collinearity. On the basis of the discussions on multi-collinearity above, for detecting the presence of multi-collinearity and estimating the parameters, we therefore applied a method based on Frisch's confluence analysis.

If the existence of multi-collinearity is found in our function, the adoption of a particular solution would be dependent on several factors, such as, availability of other sources of data for large samples, the importance of variables which are multi-collinear, the purpose for which the function is being estimated etc. However, if multi-collinearity has serious effects on.the coefficient estimates, we have different options to adopt. Among the methods, suggested in literature, we could not adopt the method of increasing the size of the sample since it is fixed in our case.We also could not use Distributed-lag Models by substituting lagged

Table 6.1

Partial Correlation of all variables for the year 1970

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1.00000						
X_2	0.74494*	1.00000					
X_3	0.87893*	0.94036*	1.00000				
X_4	0.75446*	0.72064*	0.82499*	1.00000			
X_5	0.37155	0.39360	0.27346	0.26908	1.00000		
X_6	0.28549	0.34921	0.33736	0.29104	0.08753	1.00000	
X_7	0.92908*	0.87165*	0.92638*	0.72563*	0.46444	0.26688	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.2

Partial Correlation of all variables for the year 1971

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.73099*	1.00000					
x_3	0.84064*	0.96290*	1.00000				
x_4	0.74431*	0.73728*	0.82722*	1.00000			
x_5	0.53609*	0.54105*	0.46819	0.42965	1.00000		
x_6	0.27011	0.34805	0.35030	0.29883	0.11941	1.00000	
x_7	0.90498*	0.89727*	0.92392*	0.73250*	0.65231*	0.26688	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.3

Partial Correlation of all variables for the year 1972

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.88111*	1.00000					
x_3	0.85748*	0.98886*	1.00000				
x_4	0.78121*	0.83019*	0.88089*	1.00000			
x_5	0.43480	0.30999	0.21728	0.24000	1.00000		
x_6	0.27878	0.33824	0.34545	0.32385	0.00641	1.00000	
x_7	0.91907*	0.96039*	0.92879*	0.78473*	0.45708	0.26688	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.4

Partial Correlation of all variables for the year 1973

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.88111*	1.00000					
x_3	0.85748*	0.98886*	1.00000				
x_4	0.78131*	0.83019*	0.88089*	1.00000			
x_5	0.43480	0.30999	0.21728	0.24000	1.00000		
x_6	0.27878	0.33824	0.34545	0.32385	0.00641	1.00000	
x_7	0.91907*	0.96039*	0.92879*	0.78473*	0.45708	0.26688	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.5

Partial Correlation of all variables for the year 1974

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.66613*	1.00000					
x_3	0.83912*	0.86025*	1.00000				
x_4	0.77737*	0.79180*	0.88161*	1.00000			
x_5	0.48309	0.61347*	0.39233	0.45247	1.00000		
x_6	0.28454	0.33674	0.33126	0.31982	0.11936	1.00000	
x_7	0.92500*	0.82802*	0.93350*	0.79394*	0.55957*	0.26688	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.6

Partial Correlation of all variables for the year 1975

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.81505*	1.00000					
x_3	0.93645*	0.90687*	1.00000				
x_4	0.85587*	0.80000*	0.89513*	1.00000			
x_5	0.70296*	0.81120*	0.68275*	0.65036*	1.00000		
x_6	0.30044	0.33761	0.33629	0.32269	0.16597	1.00000	
x_7	0.95637*	0.86396*	0.93482*	0.80973*	0.78376*	0.26849	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.8

Partial Correlation of all variables for the year 1977

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.86960*	1.00000					
x_3	0.96177*	0.90243*	1.00000				
x_4	0.85240*	0.76855*	0.86688*	1.00000			
x_5	0.80704*	0.93369*	0.79781*	0.66265*	1.00000		
x_6	0.30462	0.33313	0.33048	0.36434	0.24399	1.00000	
x_7	0.94936*	0.85925*	0.94154*	0.77037*	0.84031*	0.26849	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.9

Partial Correlation of all variables for the year 1978

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.88051*	1.00000					
x_3	0.95130*	0.94053*	1.00000				
x_4	0.82550*	0.77128*	0.85606*	1.00000			
x_5	0.79869*	0.86475*	0.78355*	0.65206*	1.00000		
x_6	0.32666	0.34162	0.34164	0.39735	0.28727	1.00000	
x_7	0.94819*	0.88321*	0.94364*	0.76422*	0.82488*	0.28376	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.10

Partial Correlation of all variables for the year 1979

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.89582*	1.00000					
x_3	0.96574*	0.95020*	1.00000				
x_4	0.83222*	0.76799*	0.86562*	1.00000			
x_5	0.80810*	0.84817*	0.78960*	0.64310*	1.00000		
x_6	0.32582	0.33994	0.35090	0.39833	0.27389	1.00000	
x_7	0.96003*	0.85980*	0.93297*	0.76449*	0.81162*	0.28376	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.11

Partial Correlation of all variables for the year 1980

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.79878*	1.00000					
x_3	0.95510*	0.84382*	1.00000				
x_4	0.74526*	0.58338*	0.78773*	1.00000			
x_5	0.69457*	0.89423*	0.66958*	0.46845	1.00000		
x_6	0.22809	0.21626	0.24851	0.33296	0.20327	1.00000	
x_7	0.94494*	0.81115*	0.91793*	0.63450*	0.71798*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.12

Partial Correlation of all variables for the year 1981

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.86621*	1.00000					
x_3	0.96533*	0.90826*	1.00000				
x_4	0.76477*	0.62917*	0.78482*	1.00000			
x_5	0.69680*	0.80686*	0.65850*	0.47484	1.00000		
x_6	0.24374	0.22427	0.25097	0.32512	0.22027	1.00000	
x_7	0.91927*	0.84243*	0.91838*	0.63636*	0.68874*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.13

Partial Correlation of all variables for the year 1982

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.89193*	1.00000					
x_3	0.97595*	0.91428*	1.00000				
x_4	0.78506*	0.65773*	0.79870*	1.00000			
x_5	0.66587*	0.78276*	0.61794*	0.42190	1.00000		
x_6	0.23791	0.22701	0.25194	0.31480	0.21446	1.00000	
x_7	0.92635*	0.85848*	0.91680*	0.64474*	0.67116*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.14

Partial Correlation of all variables for the year 1983

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.90414*	1.00000					
x_3	0.97167*	0.89042*	1.00000				
x_4	0.78971*	0.67285*	0.80622*	1.00000			
x_5	0.56841*	0.72447*	0.50742*	0.33312	1.00000		
x_6	0.24297	0.22820	0.25009	0.30193	0.20854	1.00000	
x_7	0.94090*	0.84388*	0.91272*	0.64697*	0.59035*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.15

Partial Correlation of all variables for the year 1984

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1.00000						
X_2	0.86619*	1.00000					
X_3	0.97555*	0.84627*	1.00000				
X_4	0.77560*	0.68147*	0.81614*	1.00000			
X_5	0.59040*	0.78204*	0.49366	0.36057	1.00000		
X_6	0.21578	0.21827	0.24914	0.29743	0.16810	1.00000	
X_7	0.93550*	0.79966*	0.90952*	0.65911*	0.57583*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.16

Partial Correlation of all variables for the year 1985

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	1.00000						
x_2	0.84934*	1.00000					
x_3	0.97625*	0.81327*	1.00000				
x_4	0.78520*	0.67725*	0.81749*	1.00000			
x_5	0.60963*	0.84058*	0.50940*	0.38545	1.00000		
x_6	0.23070	0.21230	0.25000	0.29462	0.15709	1.00000	
x_7	0.93372*	0.77555*	0.91108*	0.66029*	0.58414*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

Table 6.17

Partial Correlation of all variables for the year 1986

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1.00000						
X_2	0.86787*	1.00000					
X_3	0.97471*	0.83912*	1.00000				
X_4	0.77395*	0.68270*	0.80531*	1.00000			
X_5	0.51431*	0.72689*	0.40331	0.27159	1.00000		
X_6	0.22969	0.21940	0.24976	0.29805	0.11851	1.00000	
X_7	0.93331*	0.80150*	0.91650*	0.65302*	0.51101*	0.17368	1.00000

* mark indicates the possibility of presence of multi-collinearity.

variables for explanatory variables as suggested by Koyck, since continuity in substituting over the period of study is not possible as all states did not get credit in each year. Removal of fetters from the process of formation of simultaneous equation is not also possible as the introduction of additional meaningful relationship between multi-collinear x's is not sensible. Therefore, we are left with no options but to apply Principal Component Analysis, if multiple regression fails as a prehensile technique.

6.2.2 Frisch's Confluence Analysis or Bunch-map Analysis:

We adopt a somewhat revised version of Frisch's Confluence Analysis method[8]. Frisch's method estimates all possible regressions. It takes each of the variables of the function successively as a dependent variable and introduces others gradually as independent variable for the analysis. It thus finds a number of sets of regressions for each variable on others gradually incorporating one after another. It requires $n(n-1)$ number of computations of regressions for n variables present in the function and makes comparison more complicated than additional contribution to the analysis.

We here, at first, regress dependent variable on each one of the explanatory variable separately. We then examine the results of all these elementary regressions on the basis of *a priori* and statistical criteria. We select most plausible regression and insert remaining variables gradually to examine their effects on the coefficients (λ_i , i = 2, ,7), on their standard errors and on the overall R^2. Additional variable is classified as *useful*, *superfluous* or *detrimental*, if

i. the new variable improves R^2 without being cause of any individual coefficient unacceptable on a *priori* considerations, the variable is retained as an useful explanatory variable,

ii. the new variable neither improves R^2 nor affects considerably the values of the individual coefficient, the variable is rejected as superfluous,

iii. the new variable affects considerably the coefficient, either the signs or the values, then it is considered as *detrimental*. The multi-collinearity would be a serious problem in our case, if individual coefficients, being affected by the introduction of a new variable, become unacceptable on both *priori* or theoretical considerations.

We have already observed that x_1 is highly correlated with x_2 ,x_3 ,x_4 and x_7 individually; but x_2 also partially intercorrelated with x_3 , x_4 and x_7 with a considerable high order. Similarly partial correlations of x_3 with x_4 (r_{34}^{70} = 0.825) and x_7 (r_{47}^{70} = 0.926) are also high (Tables 6.1 - 6.17). 10 Therefore, intercorrelation of such as high order between explanatory variables may cause multi-collinearity of the equation (1).

For each year from 1970 to 1986, we first obtained the regression equation,

$$x_1 = f(x_j) \text{ for } j = 2, 3, \qquad ,7 \qquad (A)$$

and subsequently introduced other explanatory variables gradually in accordance with the variables appeared as useful, detrimental or superfluous. The results of these experiments are given in *Appendix II*

In the year 1970, from the estimated equation $\hat{x}_1 = \hat{\lambda} + \hat{\lambda}_j\, x_j$, for $j = 2,3,\ldots\ldots,7$ we find that for $j = 7$, the goodness of fit is the highest ($R^2 = 0.863$), $\hat{\lambda}_7$ is significant by S.E. test ($s(\hat{\lambda}_7) < \hat{\lambda}_7/2$) and also by t-test ($t^* > t_{0.025}$) at 5 per cent level as well as 1 per cent level ($t^* > t_{0.005}$). Other variables with descending order of their R^2 values are x_3, x_4, x_2, $x_5$and x_6. Introducing x_3 with x_7, though R^2 improves but $\hat{\lambda}$ and $\hat{\lambda}_3$ fail to exist. Both are insignificant by t-test and S.E. test. Replacing x_3 by x_4 , we get an improved R^2 (0.877) and also coefficients do not change adversely, and thus, x_4 may be accepted as an useful explanatory variable, x_3, joining with x_7, x_4,does not improve R^2 either and adversely changes $\hat{\lambda}_3$. Therefore, certainly x_3 is a detrimental variable. Similarly, adding x_2, x_5, x_6 in place of x_3, one after another, we find x_2 and x_5 as detrimental and x_6 superfluous. None other combination comes out as plausible, and hence, most likely the estimated relationship for this year could be

$$\hat{x}_1 = -3.06^{**} + 9.45^{**} x_7 + 71.47\ x_4 \qquad (2)^{\dagger}$$
$$(1.28) \quad (1.27) \quad (45.83)$$
$$(R^2 = 0.877,\ DW = 2.15)$$

Therefore, for this year, implicit relation is

$$x_1 = f(x_7,\ x_4)$$

In 1971, the elementary regression (as given A) shows that x_7, x_3, x_2, x_4 and x_5 is the order of the individual variable with respect to their extent of variation explained. Adding any variable with x_7 shows x_6 non-useful, x_3 and x_4 non detrimental in association with x_7 and x_5 as detrimental. But, like previous year, if x_3 is replaced by x_4. R^2 improves considerably from 0.819 to 0.833, intercept becomes significant by S.E. test, $\hat{\lambda}_7$ does not move in inappropriate direction and most important, very instable $\hat{\lambda}_3$ is replaced by a comparatively stable $\hat{\lambda}_4$. As x_3 also, when joins with x_7 ,x_4 gives coefficient $\hat{\lambda}_3$ which is totally unacceptable on priori grounds, we reject x_3 as detrimental. Since, gradual inclusion of any variable or combination of them further, does not give a relationship which maintains our conditionality of the analysis, we conclude the

---

$\dagger$ * and ** indicate parameter is significant at 5 per cent or 1 per cent level of significance by the t-test. Figures in brackets are the standard error of the estimates.

following estimated relationship for this year:

$$\hat{x}_1 = \underset{(1.69)}{-3.0} + \underset{(1.70)}{10.33^{**}} x_7 + \underset{(56.11)}{77.09}\, x_4 \qquad (3)$$
$$(R^2 = 0.833,\ DW = 2.32)$$

In the next year, 1972, we do not observe any remarkable change. The variables maintain their order of individual influence on x_1 and x_7 being the most important explanatory factor, any more variable when associated with it x_2 and x_5 appear as detrimental and x_6 superfluous. Out of the remaining two variables, x_3 and x_4, the former one, though not expresses detrimental effect on x_1 but the parameters $\hat{\lambda}_1$ and $\hat{\lambda}_3$, being very much instable, become non-existent. Thus only one combination of two variables is possible i.e. the combination of x_7 and x_4.

Gradually including other variables in the function, results proves that no other better estimated relationship exists as the effect of multi-collinearity is serious for other $\hat{\lambda}_j$. Thus we have the following estimated relationship:

$$\hat{x}_1 = \underset{(2.25)}{-2.09} + \underset{(2.35)}{10.81^{*}} x_7 + \underset{(71.58)}{81.15}\, x_4 \qquad (4)$$
$$(R^2 = 0.766,\ DW = 2.37)$$

We have, of course, noted that though the parameters are better than any other estimated relationship (vide Appendix II) but those are not altogether stable and insignificant except $\hat{\lambda}_7$.

We observe a change in 1973. In this year x_2 becomes important, as given by the elementary regression (A), compared to x_3 and x_4. But, x_2 does impair the estimation when join with x_7 as $\hat{\lambda}_2$ becomes statistically insignificant and totally unacceptable on priori consideration. Replacing x_2 by x_3 does not change the situation to a large extent. R^2 remains same as (0.845) and $\hat{\lambda}_3$ appears very instable. x_4 in place of x_2 or x_3, improves R^2 from 0.845 to 0.854 and coefficients do not change in inappropriate way. Therefore, it may considered somewhat useful. x_5 and x_6 prove themselves detrimental and superfluous as usual. Therefore, rejecting the variables on which the effect of multi-collinearity makes hindrance to estimate properly, we have the following estimated relationship:

$$\hat{x}_1 = -5.92 + 16.99^{**}x_7 + 92.61\ x_4 \qquad (5)$$
$$(2.56) \quad (2.80) \qquad (77.94)$$
$$(R^2 = 0.854,\ DW = 2.24)$$

Here again we should note that stability of $\hat{\lambda}_4$ is questionable.

In the year 1974, the order of individual variables in order of their contribution to the explained part of the elementary regression given in (A), is similar to that of 1972. None of the variables x_3, x_2, x_5 indicates significant or appropriate direction with regard to theoretical framework of the hypothesis, if combined with x_7. Effect of x_6 does not make any difference. Only x_4 gives comparatively better R^2(0.861) among acceptable estimated relationships. Variables x_5, x_6 again destabilise the relationships when each of them joins with x_7, x_4 ,and thus left no doubt about their detrimental nature. Continuing comparative analysis of the results given in the *Appendix II*, we conclude that $x_1 = f(x_7, x_4, x_2)$ and the estimated relationship is,

$$\hat{x}_1 = \underset{(3.31)}{-6.91^{*}} + \underset{(3.39)}{23.48^{**}}x_7 + \underset{(95.58)}{84.88}\, x_4 \qquad (6)$$

$$(R^2 = 0.861,\ DW = 1.99)$$

This year is also, like previous years, not the exception so far the test of significance of $\hat{\lambda}_4$ is concerned. Since introduction of x_4 impress R^2 and gives better coefficients values we may consider it as an explanatory variable with, of course, a note of doubt. We shall introduce the explanation of on it in the Section 6.2.3.

In the year of 1975, effect of x_7 on the dependent variable reaches at a high level ($R^2 = 0.915$). The variable x_3 also explains x_1 with high level of goodness of fit ($R^2 = 0.877$). x_4 and x_2 come next those two variables with 73.2 per cent and 66.4 per cent goodness of fit. Effect of x_6 is very nominal, less than 1 per cent and that of x_5 is average. Adding any one of the variables from x_2 to x_6 with x_7 only x_3 and x_4 appear as useful. Among others, x_2 and x_5 are distinctly detrimental and x_6 may be treated as superfluous but requires further test. Both x_3 and x_4, adding individually with x_7, improves the value of R^2 considerably, but the effect of x_4 is comparatively better. In both the cases, coefficients are significant in terms of S.E. test and t-test at 5 per cent level. Now adding x_3 with x_7, x_4 combination or x_4 with x_7, x_3 though R^2 improves marginally from 0.934 to 0.935 & 0.929 to 0.935 respectively but the coefficient $\hat{\lambda}_3$ and $\hat{\lambda}_4$ becomes insignificant by any test. Therefore, the combined effects of x_3 and x_4 cannot be considered. Introduction of either x_2 or x_5 with x_7, x_4 show detrimental effect and x_6 improve neither R^2 nor the coefficients. Therefore x_6 is a superfluous variable. Similarly, x_2 and x_5, each in association with x_7, x_4 combination proves themselves as detrimental and x_6 shows superfluous effect. All other possible combinations with higher number of variables gradually as expected from the analysis up to two variable-combination, does not show any improvement. Hence, either of the two combinations may be considered

$$\hat{x}_1 = -13.22^{**} + 28.81^{**}x_7 + 191.09^{**}x_4 \qquad (7)$$
$$(2.60) \qquad (3.08) \qquad (73.78)$$
$$(R^2 = 0.934,\ DW = 2.206)$$

$$\hat{x}_1 = -5.77 + 21.67^{**}x_7 + 0.015^{*}x_3 \qquad (8)$$
$$(3.84) \qquad (5.28) \qquad (0.007^{*})$$
$$(R^2 = 0.929,\ DW = 2.023)$$

Intercept in (8) is statistically insignificant. R^2 is less than the estimate of (7) by 0.5 per cent, the most plausible functional relation should be,

either $x_1 = f(x_7, x_4)$ or $x_1 = f(x_7, x_3)$

In the year 1976, the initial estimate of $x_1 = f(x_j)$; $j = 2,3,\ldots 7$ gives the order of the variables with respect to the variation explained is same as 1975 but the x_7 and x_3 becomes more close with the value of R^2 as 0.914 and 0.906 respectively.

Individually x_6 appeared, as before, with less than 1 per cent capacity of explanation and also with insignificant coefficients and intercept. Again, x_5 shows the detrimental effect, x_6 and x_2 show superfluous effect and x_3 and x_4 useful effect when combined with x_7 individually. With the elimination process through the comparative analysis as before, we have the following two most plausible relationships,

$$\hat{x}_1 = -5.83 + 24.10^{**}x_7 + 0.02^{**}x_3 \qquad (9)$$
$$(5.00) \quad (6.68) \quad (0.006)$$
$$(R^2 = 0.940, \; DW = 2.01)$$

$$\hat{x}_1 = -20.72^{**} + 32.68^{**}x_7 + 289.55^{**}x_4 \qquad (10)$$
$$(3.23) \quad (3.75) \quad (78.96)$$
$$(R^2 = 0.946, \; DW = 2.01)$$

For the same reasons, as in the case of 1975 estimated equation (10) is somewhat better than the equation (9) but it only differs by 0.6 per cent in R^2 value. Therefore, the desirable relationship should be,

$$\text{either } x_1 = f(x_7, x_4) \text{ or } x_1 = f(x_7, x_3)$$

In 1977, for the first-time, we observed from the single-variable function $x_1 = f(x_j)$; $j = 2, \ldots, 7$ that x_3 gives better explanation than x_7 and x_5 alone gives an explanation more than 65 per cent. In 1970, it was only 13 per cent and its extent of distribution was around 20 per cent up to 1974. It was 1975 when it started to contribute a considerable amount like 49.4 per cent, in the next year 62.8 per cent and finally in 1979 it is 65.1 per cent. x_3 , x_2 and x_5 , being the variables of same group, the effect of multi-collinearity should have been of higher order in this year. The results of this year is very important in the sense that the combined effect of x_3 and x_4 along with x_7 makes the equation theoretically untenable with all the coefficients $\hat{\lambda}_7, \hat{\lambda}_3, \hat{\lambda}_4$ negative and insignificant. Moreover, the extent of variation explained is reduced to 3.9 per cent instead of around 90 per cent that the observed in the previous years. Secondly, (x_7 , x_3) as a combination of explanatory variables (R^2 = 0.942) is better than (x_7, x_4) combination (R^2 = 0.937). Thirdly, association of any explanatory variable with either the combination of (x_7 , x_4) or (x_7 , x_3) is proved as detrimental with distinctly lower value of R^2 except in the case of the addition of x_2 with the former. If x_2 is added, R^2 remains same as 0.942 but the coefficients $\hat{\lambda}_2$ unacceptable with the theoretical value as well as insignificant by the S.E. test and t-test. Therefore, we have,

$$\hat{x}_1 = \underset{(6.37)}{-5.04} + \underset{(8.78)}{22.73^{**}}x_7 + \underset{(0.007)}{0.03^{**}}x_3 \qquad (11)$$
$$(R^2 = 0.943,\ DW = 2.07)$$

$$\hat{x}_1 = \underset{(4.53)}{-29.22^{**}} + \underset{(4.82)}{42.39^{**}}x_7 + \underset{(91.72)}{333.56^{**}}x_4 \qquad (12)$$
$$(R^2 = 0.937,\ DW = 1.80)$$

R being higher only by 0.5 per cent, we should consider the following relationships as desirable,

$$x_1 = f(x_7, x_3) \quad \text{or} \quad x_1 = f(x_7,\ x_4)$$

In 1978, variables maintain the same order, as 1977, in accordance with their contribution to the extent of explained variation individually. This year also maintains the position that when the variable x_7 is combined with x_3 or with x_4 gives acceptable estimation only and the values of R^2, being 0.928 and 0.923 respectively, show (x_7 , x_3) combination more attractive than the other one. But the combination (x_7, x_3, x_4)does not fetch a miserable R^2-value like previous year. The value of R^2 , though, increases (R^2 = 0.932) but $\hat{\lambda}_3$ and $\hat{\lambda}_4$, being insignificant by S.E. and t-tests, become non-existent. Contribution of other variables in the estimation has been found either superfluous or detrimental and, as before, we have the following two estimations:

$$\hat{x}_1 = -8.85 + 32.32^{**}x_7 + 0.02^{**}x_3 \qquad (13)$$
$$(8.40) \quad (11.58) \qquad (0.007)$$
$$(R^2 = 0.928,\ DW = 2.03)$$

$$\hat{x}_1 = -34.11^{**} + 53.48^{**}x_7 + 312.04^{**}x_4 \qquad (14)$$
$$(5.80) \qquad (6.13) \qquad (112.62)$$
$$(R^2 = 0.923,\ DW = 1.90)$$

Finally due to higher R^2 though nominal amount of 0.005, either

$x_1 = f(x_7, x_3)$ or $x_1 = f(x_7, x_4)$

In 1979, we observe comparatively important position of the variable x ,then previous year, as an explanatory one, and for the first time we find that $x_1 = f(x_7, x_2)$ gives statistically acceptable estimation along with $x_1 = f(x_7, x_3)$ and $x_1 = f(x_7, x_4)$. In all the three estimations coefficients and intercepts are significant by both S.E. and t-test except (x_7, x_3) combination where intercept still remain insignificant. Addition of x_2 with (x_3 , x_7) shows x_2 as detrimental explanatory variable and with (x_7 , x_4) shows acceptable estimation with all parameters, tested through S.E. and t-test, significant. But R^2 (0.953) being less than the value of R^2 (0.959) with only two variables (x_7 , x_3), the estimation fails to be an encouraging one. Of course the value of this R^2 is an improved one in comparison with (x_7, x_4) combination (R^2 = 0.945). As other combination of explanatory variables appears either detrimental and non-useful, we conclude the following two estimation as proper and acceptable:

$$\hat{x}_1 = -35.84^{**} + 59.79^{**}x_7 + 0.006x_2 + 310.81^{**}x_4 \qquad (15)$$
$$(8.97) \quad (8.51) \quad (0.002) \quad (124.20)$$
$$(R^2 = 0.953,\ DW = 1.66)$$

$$\hat{x}_1 = -14.29^{**} + 41.66^{**}x_7 + 0.03^{**}x_3 \qquad (16)$$
$$(7.82) \quad (10.43) \quad (0.006)$$
$$(R^2 = 0.959,\ DW = 1.76)$$

Therefore, estimation, if expressed in implicit function, can be either

$$x_1 = f(x_7,\ x_2,\ x_4) \quad \text{or} \quad x_1 = f(x_7,\ x_3)$$

Dependence of the variable x_1 on x_2 is increased in this year (Table 6.10) and also the interdependence between x_2 and x_3. But, at the same time, inter dependencies between x_2,x_7 and x_2 ,x_4, one decreased with respect to previous years. This is possibly the reason that x_2 could not combine with x_3 and x_7 due to the presence of high multi-collinearity and x_3 and x_7 highly correlated with x_1 ($r_{x1,x3} = 0.966$, $r_{x1,x7} = 0.960$). But due to lower degree of dependency of x_1 on x_4 ($r_{x1,x4} = 0.832$) and also lower value of partial correlation coefficient between x_4, x_2, make x_4, x_2 together to combine with x_7 ($r_{x4,x2} = 0.767$).

In the year 1980, effects of x_4 as well as x_2 on $x_4$are much less than the previous year (Table 6.11). Partial correlation between x_2 and x_4 also comes down substantially to 0.583 from 0.767 as in the previous year. x_2 and x_7 is also correlated with lower extent (0.811) than previous year

(0.859). Thus, the resultant effect of all these changes on the estimation could be interesting for the analysis. Single variable estimations through (A) do not change the preference of variables for inclusion in the regression equation. After identifying the variables as useful, detrimental or superfluous from the results we obtain,

$$\hat{x}_1 = -14.72 + 42.99^{**}x_7 + 0.032^{**}x_3 \qquad (17)$$
$$(9.06) \quad (12.31) \qquad (0.007)$$
$$(R^2 = 0.941,\ DW = 1.38)$$

$$\hat{x}_1 = -55.68^{**} + 78.37^{**}x_7 + 462.69^{**}x_4 + u_2 \qquad (18)$$
$$(8.06) \quad (7.007) \qquad (134.04)$$
$$(R^2 = 0.928,\ DW = 1.46)$$

R^2 being substantially high, we hold that the possible estimation, stated by implicit function, should be

$$x_1 = f(x_7, x_3)$$

Results of 1981 explains the situation different to that of the year 1979. In the final analysis, variables after being rejected as non-useful or detrimental, we isolate the following variables with the relation given below:

$$\hat{x}_1 = -0.29 + 25.90x_7 + 0.04^{**}x_3 \qquad (19)$$
$$(11.81) \quad (15.83) \quad (0.007)$$
$$(R^2 = 0.938)$$

$$\hat{x}_1 = -52.04^{**} + 68.85^{**}x_7 + 0.009x_2 + 631.12^{**}x_4 \qquad (20)$$
$$\quad (15.91) \quad (14.54) \quad (0.004) \quad (192.89)$$
$$(R^2 = 0.914,\ DW = 1.64)$$

The estimation (20) is acceptable with all parameters tested as significant, but R^2 is substantially higher in (19). However the intercept and $\hat{\lambda}_7$, when x_3 joins with x_7, fails to appear as significant by S.E test and by the t-test at 5 per cent level of significance. Whereas, x_2 combining with x_7 and x_4 fetches an improved $R^2 = 0.899$ and $\hat{\lambda}_7$, $\hat{\lambda}_2$, $\hat{\lambda}_4$ are confirmed by S.E. and t-test except $\hat{\lambda}_2$ which, though evolves as significant by S.E. test does not pass test of significance by t-test at 5 per cent level. Therefore, the equation (19) becomes only appropriate relationship, as the relation (18), having higher R^2 , does not exist due to insignificant intercept and coefficient.

Scenario in 1982, though single variable OLS estimations given by (A), maintains the order of preference of explanatory variables as before, but when one more variable is added (x_2 , x_7) combination also comes out as acceptable along with (x_7 , x_4) and R^2 values look very satisfactory for each of the estimation. Of course R^2 is the highest for x_7 , x_3 combination and that also substantially differs from above two (0.959 as against 0.893 and 0.918).

But test of significance rules out x_7, x_3 combination as a set of explanatory variables as both intercept and coefficients become insignificant. As the introduction of x_3 with x_7 makes $\hat{\lambda}_7$ and the intercept unacceptable by t-test and S.E. test, we can not consider it. x_7, x_4 combination looks more attractive due to higher R^2 among two and significant parameters. When x_2 is introduced with it, R^2 improves to 0.935 and parameters remain acceptable. Since introduction of other variables does not give better explanation, we hold that, for the year 1982, estimated equation as,

$$\hat{x}_1 = -63.22^{**} + 74.99^{**}x_7 + 0.011^{**}x_2 + 749.25^{**}x_4 \qquad (21)$$
$$(16.88) \quad (15.10) \quad (0.004\ ) \quad (196.62)$$
$$(R^2 = 0.934,\ DW = 1.69)$$

(x_7, x_2), (x_7, x_3) and (x_7, x_4) combination in 1983 gives acceptable estimations among which (x_7, x_3) gives the best the best explanation of the variation ($R^2 = 0.961$). parameters are also significant by S.E. and t-test at 5 per cent level of significance in all three cases. (x_7, x_3, x_4) and (x_7, x_2, x_3), each as a set of explanatory variables is not acceptable since $\hat{\lambda}_4$ becomes insignificant by both the tests of significance in the first set and $\hat{\lambda}_2$ in the second set. The only OLS estimation with these explanatory variables which seems to be appreciable due to the enhanced R^2 value from 0.941 to 0.961 and parameters significant, is with (x_7, x_2, x_4). Inclusion of any variable within the set does violate our proposed scheme of acceptability except x_3 variable. When this variable is also included those, jointly explain the variation in a better way as expressed by R^2 value of 0.971. Other variables seem to be either unacceptable as detrimental or non-useful as superfluous. Therefore, the only possible estimation becomes

$$\hat{x}_1 = -47.17^{*} + 65.49^{**}x_7 + 0.008x_2 + 0.026^{**}x_3 + 451.52x_4 \qquad (22)$$
$$(20.78) \quad (17.37) \quad (0.004) \quad (0.009) \quad (220.06)$$
$$(R^2 = 0.971)$$

In this relation (22) we see that $\hat{\lambda}_2$ and $\hat{\lambda}_4$ though do not satisfy t-test at 5 per cent level, but it is acceptable by S.E. test.

OLS estimation in 1984 as mentioned in (A) place the variables in similar order of preference in accordance with their extent of variation explained and also (x_7, x_3) combination explained to the highest extent (R^2 = 0.965) when combined any two variables. $\hat{\lambda}_7$ and $\hat{\lambda}_3$ also exist with significance by S.E. test and by t-test, even at 1 per cent level. So, this is the best estimated regression on x_1. Adding x_4, and x_6 with this, one after another, does not show any improvement but additional effect of x_2 and x_5 separately with the former gives better estimation. But contribution of x_5 is far better than x_3 (R^2 = 0.973 and 0.969 respectively). Also $\hat{\lambda}_1$ is insignificant. x_7, x_2, x_4 jointly give an acceptable estimation, though R^2 is lower than the former (0.937). We, therefore, conclude the following two estimations as acceptable,

$$\hat{x}_1 = -100.37^{**} + 127.15^{**} x_7 + 0.01^{**} x_2 + 871.61^{**} x_4 \qquad (23)$$
$$(24.99) \quad (18.92) \quad (0.004) \quad (296.42)$$
$$(R^2 = 0.937)$$

$$\hat{x}_1 = -41.03^{**} + 41.58^{**} x_7 + 0.05^{**} x_3 + 63.56^{**} x_5 \qquad (24)$$
$$(15.32) \quad (18.41) \quad (0.005) \quad (24.03)$$
$$(R^2 = 0.973)$$

Though both are acceptable, (24) gives better estimation. Thus, for the first time, variable x_5 assumes important position as an explanatory factor.

In 1985, we get the following three regression equations as acceptable:

$$\hat{x}_1 = -17.34 + 59.33^{**} x_7 + 0.006^{**} x_2 + 0.047^{**} x_3 \qquad (25)$$
$$(16.59) \quad (22.22) \quad (0.002) \quad (0.006)$$
$$(R^2 = 0.971)$$

$$\hat{x}_1 = -144.837^{**} + 155.98^{**} x_7 + 0.009^{**} x_2 + 1197.35^{**} x_4 \qquad (26)$$
$$(31.56) \quad (21.85) \quad (0.003) \quad (372.63)$$
$$(R^2 = 0.939)$$

$$\hat{x}_1 = -52.60^{**} + 44.28^{**} x_7 + 0.05^{**} x_3 + 76.71^{**} x_5 \qquad (27)$$
$$(17.16) \quad (21.89) \quad (0.005) \quad (25.46)$$
$$(R^2 = 0.975)$$

x_7 being common, either x_2 or x_5 can give meaningful estimation when combined with x_3 , and any of these combination results in higher R^2 than joint explanation of (x_2, x_4). Moreover, equations (24) and (26) differs only by 0.4 per cent variation and therefore, both of them can be considered as plausible. Hence,

either $x_1 = f(x_7, x_2, x_3)$ or $x_1 = f(x_7, x_3, x_5)$

would be equally likely implicit relations. But it is to be noted that adding x_5 with (x_7, x_3) makes $\hat{\lambda}_7$ insignificant so far t-test is concerned.

In a similar way, in 1986, we obtain only two estimations as (x_7, x_3, x_5) together fails to be acceptable due to insignificant $\hat{\lambda}_7$ by both the tests. Therefore, we have, for this year, the following acceptable estimation:

$$\hat{x}_1 = \underset{(20.04)}{-11.29} + \underset{(27.14)}{61.67^{**}}x_7 + \underset{(0.003)}{0.007}x_2 + \underset{(0.007)}{0.04^{**}}x_3 \qquad (28)$$
$$(R^2 = 0.966)$$

$$\hat{x}_1 = \underset{(36.40)}{-149.06^{**}} + \underset{(25.47)}{167.63^{**}}x_7 + \underset{(0.005)}{0.01^{**}}x_2 + \underset{(428.93)}{1272.16^{**}}x_4 \qquad (29)$$
$$(R^2 = 0.936)$$

Hence, R^2 being substantially higher by 3 per cent, it seems that the acceptable function would be,

$$x_1 = f(x_7, x_2, x_3)$$

But, it is to be noted that $\hat{\lambda}_2$ is only significant by S.E. test not by t-test.

6.3 Principal Component Analysis and its Results:

From the above discussions of the results obtained from Frisch's confluence analysis, we observed that due to multi-collinearity, we could not estimate our functions with all explanatory variables. But to reject all explanatory variables except x_7 seriously impaired our basic objective of estimation. We, therefore, resort to application of Principal Component Analysis as we have mentioned earlier.

The results of the principal component analysis have given in Tables 6.18 -6.34. All principal components are extracted and presented in the Tables for comparisons.

Then we used Kaiser's criterion for selecting the number of principal components, maximum number of which may be equal to 6 in our case. For estimating the level of significance of loadings, we used the test based on Pearson product moment correlation coefficients for different sample sizes.

One may argue for Burt-Banks tests for better estimation as it considers both the number of variables and the order of extractions of the principal components, whereas the test used by us considers neither.

Burt-Banks test adjusts the standard error of the correlation

Table 6.18

Principal Components, Corresponding Loading and Latent Roots for the Year 1970

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.83	63.87	0.94	0.96	0.85	0.47^*	0.42^*	0.94
P_2^*	0.94	15.61	0.00	0.08	0.07	−0.64	0.70	−0.13
P_3^*	0.78	13.05	−0.07	−0.21	−0.21	0.59	0.57	−0.08
P_4^*	0.31	5.24	−0.21	−0.09	0.48	0.08	0.01	−0.16

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.19

Principal Components, Corresponding Loading and Latent Roots
for the Year 1971

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.06	67.60	0.95	0.96	0.85	0.65	0.41*	0.95
P_2^*	0.92	15.33	-0.01	0.03	0.03	-0.40	0.86	-0.15
P_3^*	0.62	10.36	-0.11	-0.22	-0.26	0.63	0.31	0.00
P_4^*	0.30	5.07	-0.22	-0.11	0.46	0.10	0.01	-0.15
P_5	0.09	1.47	0.19	0.01	0.00	0.05	-0.02	-0.23

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.20

Principal Components, Corresponding Loading and Latent Roots for the Year 1972

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.06	67.73	0.94	0.96	0.88	0.67	0.42^*	0.93
P_2^*	0.88	14.71	-0.04	-0.01	-0.01	-0.28	0.89	-0.14
P_3^*	0.63	10.45	-0.02	-0.25	-0.24	0.68	0.19	-0.08
P_4^*	0.25	4.18	-0.16	-0.10	0.41	0.09	-0.01	-0.19

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.21

Principal Components, Corresponding Loading and Latent Roots for the Year 1973

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.96	65.93	0.98	0.97	0.90	0.40*	0.41*	0.96
P_2	1.03	17.14	-0.01*	-0.10*	-0.09*	0.76	-0.64	0.15*
P_3^*	0.75	12.52	-0.12	-0.18	-0.14	0.51	0.65	-0.05
P_4^*	0.23	3.85	-0.15	-0.06	0.41	0.06	-0.02	-0.18

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.22

Principal Components, Corresponding Loading and Latent Roots for the Year 1974

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.03	67.15	0.93	0.94	0.90	0.64	0.41*	0.94
P_2^*	0.91	15.21	-0.06	0.04	0.02	-0.40	0.86	-0.11
P_3^*	0.67	11.13	0.04	-0.30	-0.21	0.65	0.31	-0.12
P_4^*	0.21	3.45	-0.06	-0.08	0.37	0.05	-0.02	-0.24

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.23

Principal Components, Corresponding Loading and Latent Roots for the Year 1975

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.39	73.12	0.95	0.96	0.90	0.83	0.39^*	0.95
P_2^*	0.90	15.07	-0.04	-0.01	0.00	-0.24	0.91	-0.12
P_3^*	0.39	6.48	0.09	-0.23	-0.30	0.47	0.12	-0.04
P_4^*	0.18	2.92	-0.09	-0.12	0.31	0.14	0.00	-0.20

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.24

Principal Components, Corresponding Loading and Latent Roots
for the Year 1976

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.55	75.79	0.96	0.97	0.90	0.90	0.40*	0.95
P_2^*	0.89	14.78	-0.06	-0.05	0.02	-0.18	0.91	-0.13
P_3^*	0.31	5.24	0.10	-0.13	-0.39	0.35	0.08	0.04
P_4^*	0.13	2.23	0.12	-0.09	0.14	0.12	-0.02	-0.28

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.25

Principal Components, Corresponding Loading and Latent Roots for the Year 1977

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.48	74.66	0.96	0.97	0.87	0.90	0.41*	0.94
P_2^*	0.88	14.73	-0.08	-0.06	0.06	-0.18	0.91	-0.14
P_3^*	0.37	6.21	0.18	-0.15	-0.43	0.35	0.10	-0.02
P_4^*	0.18	2.94	0.12	-0.13	0.21	0.13	-0.03	-0.29

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.26

Principal Components, Corresponding Loading and Latent Roots for the Year 1978

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.48	74.64	0.96	0.97	0.87	0.88	0.44*	0.94
P_2^*	0.86	14.41	-0.10	-0.09	0.07	-0.14	0.89	-0.16
P_3^*	0.36	5.98	0.10	-0.12	-0.42	0.38	0.09	0.02
P_4^*	0.18	2.92	-0.05	-0.15	0.23	0.23	-0.04	-0.20

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.27

Principal Components, Corresponding Loading and Latent Roots for the Year 1979

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.46	74.32	0.95	0.98	0.88	0.87	0.44*	0.93
P_2^*	0.87	14.43	-0.10	-0.08	0.07	-0.16	0.89	-0.16
P_3^*	0.36	6.00	0.09	-0.11	-0.42	0.39	0.11	0.01
P_4^*	0.17	2.83	-0.05	-0.12	0.22	0.23	-0.03	-0.22

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.28

Principal Components, Corresponding Loading and Latent Roots for the Year 1980

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.04	67.35	0.93	0.95	0.78	0.84	0.34	0.92
P_2^*	0.96	16.06	-0.16	-0.04	0.22	-0.18	0.91	-0.16
P_3^*	0.60	9.98	0.24	-0.22	-0.50	0.44	0.23	-0.08
P_4^*	0.27	4.45	0.06	-0.15	0.29	0.21	-0.08	-0.32

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.29

Principal Components, Corresponding Loading and Latent Roots
for the Year 1981

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.06	67.59	0.94	0.96	0.79	0.81	0.34*	0.92
P_2^*	0.95	15.78	-0.14	-0.07	0.17	-0.12	0.92	-0.18
P_3^*	0.56	9.26	0.15	-0.17	-0.49	0.49	0.15	-0.03
P_4^*	0.29	4.79	-0.06	-0.15	0.31	0.29	-0.09	-0.27

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.30

Principal Components, Corresponding Loading and Latent Roots
for the Year 1982

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	4.03	67.24	0.95	0.96	0.80	0.78	0.34*	0.92
P_2^*	0.94	15.71	-0.13	-0.07	0.15	-0.12	0.92	-0.18
P_3^*	0.61	10.15	0.11	-0.19	-0.48	0.55	0.15	-0.02
P_4^*	0.27	4.50	-0.05	-0.13	0.32	0.26	-0.09	-0.27

* marks over P indicates inessential as per Kaiser's Criterion, and

Table 6.31

Principal Components, Corresponding Loading and Latent Roots for the Year 1983

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.91	65.13	0.95	0.95	0.80	0.71	0.35*	0.92
P_2^*	0.93	15.57	-0.12	-0.07	0.12	-0.10	0.93	-0.18
P_3^*	0.72	12.05	0.11	-0.22	-0.45	0.66	0.12	-0.05
P_4^*	0.28	4.66	-0.02	-0.12	0.36	0.20	-0.09	-0.29

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.32

Principal Components, Corresponding Loading and Latent Roots for the Year 1984

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.89	64.82	0.94	0.94	0.82	0.72	0.33^*	0.91
P_2^*	0.95	15.76	-0.14	-0.02	0.15	-0.22	0.91	-0.14
P_3^*	0.72	12.00	0.17	-0.26	-0.40	0.64	0.21	-0.12
P_4^*	0.30	4.92	0.06	-0.12	0.37	0.12	-0.08	-0.34

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.33

Principal Components, Corresponding Loading and Latent Roots for the Year 1985

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.90	64.97	0.94	0.94	0.82	0.74	0.33*	0.90
P_2^*	0.95	15.89	-0.16	0.00	0.16	-0.25	0.91	-0.12
P_3^*	0.72	11.94	0.21	-0.29	-0.38	0.60	0.25	-0.16
P_4^*	0.31	5.13	0.09	-0.12	0.38	0.08	-0.08	-0.36
P_5^*	0.09	1.53	-0.20	-0.10	0.10	0.12	0.00	0.12

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

Table 6.34

Principal Components, Corresponding Loading and Latent Roots for the Year 1986

Principal Components (P_i)	Latent Root (λ_i)	Variation Explained	Loadings					
			l_{i2}	l_{i3}	l_{i4}	l_{i5}	l_{i6}	l_{i7}
P_1	3.78	63.05	0.94	0.94	0.81	0.64	0.33^*	0.91
P_2^*	0.98	16.27	-0.17	0.05	0.24	-0.41	0.84	-0.11
P_3^*	0.79	13.13	0.13	-0.27	-0.34	0.63	0.42	-0.15
P_4^*	0.31	5.10	0.07	-0.13	0.39	0.11	-0.08	-0.33
P_5^*	0.11	1.89	-0.25	-0.06	0.11	0.12	0.00	0.14

* marks over P indicates inessential as per Kaiser's Criterion, and

* marks over l_{ij} indicates loading is non-significant at 1 percent level as per Pearson's Critical value in the case of essential principal component.

coefficients, that we used for estimation, in order to obtain the standard error of the loadings by the following relation

$$s(l_{mj}) = \{s(r_{x_i x_j})\} \sqrt{\frac{k}{k+1-m}}$$

where, k = number of X 's in the set

m = the order of extraction of P

From Tables 6.18 - 6.34 we find that only the first component in each year came out as essential by applying Kaiser's criterion that the component P_m is to be retained if eigen value $\lambda_m > 1$. There is logical, theoretical and empirical support for this criterion[9]. Result of the year 1973 is only exception where first two components appeared as essential. All non-essential components are marked by * in our Tables.

Therefore, except 1973, for all the years we have m = 1. The Burt-Banks adjustment relation, given above, shows that both the tests are identical if m = 1.

Loadings of the variables x_6 for all the years from 1970 to 1986 are appeared as linearly insignificant at 1 per cent level. All other variables' loadings are significant at 1 per cent level of significance throughout the period of our study barring the variable x_5, loading of which is insignificant for then year 1970 and 1973. Though we have found only one orthogonal component for all the years, except 1973, this component absorbs much of the variations. The absorption is to the extent of minimum 63 per cent and in some years, namely 1975 to 1979, this is above 73 per cent. Therefore, if we consider the single orthogonal component along with the variables, loadings of which are significant, then we would not lose much of the information. Hence the component that we have extracted can be expressed as,

$$P_1 = \sum_{j=2}^{7} l_{ij} \; z_j \qquad (30)$$

where z = the standardised values of the original x's.

Therefore, if one regresses dependent variable x_1 on the extracted principal component, in our case, it would be simply,

$$x_1 = \gamma + \gamma_1 P_1 + v \qquad (31)$$

or,

$$x_1 = \gamma + \gamma_1 P_1 + \gamma_2 P_2 + v \qquad (31A)$$

where v is a random variable satisfying the usual assumptions.

From equation (4), therefore, we can obtain OLS estimates $\hat{\gamma}_1$ only. Obtaining $\hat{\gamma}_1$, we find that the original coefficients have the following reduced form of relationships:

$$\left.\begin{array}{l}\hat{\gamma}_1\ \hat{a}_{21} = \hat{\lambda}_2 \\ \hat{\gamma}_1\ \hat{a}_{31} = \hat{\lambda}_3 \\ \hat{\gamma}_1\ \hat{a}_{41} = \hat{\lambda}_4 \\ \hat{\gamma}_1\ \hat{a}_{51} = \hat{\lambda}_5 \\ \hat{\gamma}_1\ \hat{a}_{61} = \hat{\lambda}_6 \\ \hat{\gamma}_1\ \hat{a}_{71} = \hat{\lambda}_7\end{array}\right] \qquad (32)$$

for 1973, it would be,

$$\hat{\gamma}_1\ \hat{a}_{j1} + \hat{\gamma}_2\ \hat{a}_{j2} = \hat{\lambda}_j \quad \text{for } j=2,3,\ldots\ldots,7 \quad \ldots(33)$$

Results of the regression (31) or (31A) are given in the Table 6.35. Value of $\hat{\lambda}_2$, applicable only for the year 1973, is not shown in the Table 6.35. Stability of $\hat{\lambda}$ and $\hat{\lambda}_1$, as it should be, is confirmed by the fact that, for all the years from 1970 to 1986, those are significant at 1 per cent level by the t-test. We have thus calculated $\hat{\lambda}_j; j=2,3,\ldots,7$ by the relations given by (31) or (31A). These values of $\hat{\lambda}_j$ are given in Table 6.36.

6.4 Summary of the results:

From Tables 6.35 and 6.36, we find that contribution of x_7 to the value of x_1 remains important throughout the period of our study. Similarly, x_2 and x_3 are also very important in explaining the variation of x_1. Regarding x_6, we find that since $\hat{\lambda}_6$ never assumes such a value which can be as recognised as substantial. During the initial phase after nationalisation, x_4 influences the SSI credit as an important determining factor compare to x_5 and x_6. This initial phase lasted at least up to 1976. Then after we find that x_5 gradually started to be comparatively important than x_4. If we consider $\hat{\lambda}_4$ during this period of our study we observe that there are three distinct phases.

First, from 1970 to 1976 when x_4 were relatively more important than x_5, Second; from 1977 to 1981 during which it contributed comparatively less than x_5, and Third; the remaining period of our study, when it regained its earlier position. The results are complete agreement with our findings and discussion in 6.2.2. In that sub-section we find that x_6 was superfluous which, in fact, is being supported here by the values of the coefficients $\hat{\lambda}_6$ obtained from principal component Analysis. It implies that spread of education in a state, in no way influences the SSI credit after nationalisation. Then how a government can justify the success of the schemes of nationalised banks to provide finance to the educated unemployed for self employment?

Second point is that, in Chapter IV and Chapter V, we observed that SSI credit distribution was tilted towards more developed region, namely Northern and Western region, and the traditionally backward regions, specially Eastern and North-Eastern region, were deprived of getting there due share. What we find in this Chapter, either by multiple regression analysis in 6.2 or by principal component analysis in 6.3, that the variable x_7 comes out as one of the most important determining factors. x_7 , being the urbanisation index, represents the degree of development. Thus, the result obtained in earlier chapter and the contention that the inequality in distribution of credit did not go away with the nationalisation of banks, is supported by these economic techniques. Moreover, monotonically increasing value of $\hat{\lambda}_7$ (Table 6.19) also implies that the concentration of SSI credit in developed region increased over time. This, therefore, corroborates our conclusion that the inequality not only exists but also increasing in nature. In Chapter IV we observed highly significant regression coefficient of PCC on PCD. From the multiple regression analysis we get that x_2 and x_3 analysis influence x_1 to large extent and most of the years either coefficients of x_2, x_3 individually or jointly estimated as stable and significant. $\hat{\lambda}_2$ and $\hat{\lambda}_3$ of principal component analysis also show (Table 6.36) the dominence of weightages of these two variables.

In economic sense, it signifies that the level of deposit in a region mainly determines the level of credit in that region. This phenomenon,in economic parlance,indicative of credit worthiness. It is therefore, confirmation of the observation in Chapter IV,that the benefit of banking system even after nationalisation is largely restricted to a group who are likely to be both creditors and,at the same time,depositors.Thus one of the basic objectives of banks' nationalisation that the usual practice of any banking system aimed at maximising profit, 'deposits give credit worthiness' should be changed to 'credit creates deposits', has been frustrated. Results of 6.2 and 6.3 confirmed strongly this observation.

Table 6.35

Estimation of the Principal Components*

Year	$\hat{\gamma}$	$\hat{\gamma}_1$	R^2	DW
1970	5.528	0.013	0.68275	2.28
	(1.332, 4.146)	(0.002, 7.035)		
1971	6.752	0.011	0.0.62616	2.28
	(1.689, 4.145)	(0.001, 6.207)		
1972	8.626	0.011	0.49497	2.38
	(2.132, 4.046)	(0.002, 4.748)		
1973	8.609	0.020	0.75675	1.95
	(2.218, 3.880)	(0.002, 8.459)		
1974	14.084	0.018	0.059634	1.81
	(3.662, 3.888)	(0.003, 5.829)		
1975	13.157	0.015	0.77383	2.17
	(3.011, 4.370)	(0.001, 9.062)		
1976	17.737	0.013	0.80194	2.13
	(3.719, 4.770)	(0.001, 9.857)		
1977	20.117	0.015	0.84090	2.11
	(4.286, 4.694)	(0.001, 11.263)		
1978	21.893	0.017	0.84509	2.09
	(4.903, 4.466)	(0.001, 11.678)		

Table 6.35 (Contd.)

Estimation of the Principal Components

Year	$\hat{\gamma}$	$\hat{\gamma}_1$	R^2	DW
1979	25.909	0.018	0.87056	2.20
	(5.808, 4.461)	(0.001, 12.967)		
1980	30.870	0.019	0.78264	2.11
	(7.124, 4.333)	(0.002, 9.488)		
1981	33.214	0.024	0.85971	2.02
	(7.292, 4.555)	(0.001, 12.377)		
1982	32.323	0.026	0.89783	2.01
	(7.250, 4.458)	(0.001, 14.822)		
1983	35.733	0.031	0.92139	1.65
	(8.136, 4.392)	(0.001, 17.118)		
1984	48.458	0.029	0.89435	2.47
	(10.878, 4.455)	(0.002, 14.547)		
1985	56.108	0.028	0.88128	2.37
	(13.823, 4.059)	(0.002, 13.623)		
1986	54.528	0.031	0.91019	2.29
	(13.459, 4.025)	(0.001, 15.917)		

Figures within the braces indicate S.E. and t-values of the estimate.

* Value of $\hat{\gamma}_2$ which is valid for 1973 only is not shown in the table.

Table 6.36

Coefficients of original variables (x's) from the principal components estimate ($\hat{\gamma}_1$)

Year	$\hat{\gamma}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$	$\hat{\lambda}_5$	$\hat{\lambda}_6$	$\hat{\lambda}_7$
1970	0.01317	0.0124	0.0126	0.0112	0.0062	0.0055	0.0124
1971	0.01168	0.0111	0.0112	0.0099	0.0076	0.0047	0.0111
1972	0.01149	0.0108	0.0110	0.0101	0.0077	0.0048	0.0107
1973	0.02000	0.0196	0.0194	0.0180	00.80	0.0082	0.0192
1974	0.01761	0.0164	0.0165	0.0158	0.0113	0.0072	0.0164
1975	0.01596	0.0152	0.0153	0.0144	0.0132	0.0065	0.0152
1976	0.01315	0.0126	0.0127	0.0118	0.0118	0.0051	0.0125
1977	0.01511	0.0145	0.0146	0.0131	0.0136	0.0060	0.0142
1978	0.01669	0.0160	0.0162	0.0145	0.0147	0.0068	0.0157
1979	0.01822	0.0173	0.0178	0.0160	0.0158	0.0080	0.0169
1980	0.01925	0.0179	0.0183	0.0150	0.0162	0.0084	0.0177
1981	0.02365	0.0222	0.0227	0.0187	0.0192	0.0084	0.0218
1982	0.02594	0.0246	0.0249	0.0207	0.0202	0.0088	0.0239
1983	0.03115	0.0296	0.0296	0.0249	0.0221	0.0109	0.0287
1984	0.02908	0.0273	0.0273	0.0238	0.0209	0.0095	0.0265
1985	0.0.02825	0.0266	0.0266	0.0232	0.0209	0.0093	0.0254
1986	0.03117	0.0293	0.0293	0.0252	0.0199	0.0102	0.0284

References

1. Koutsoyiannis, A. Theory of Econometrics. New York,Macmillan,1987,p 233.

2. Fox,K. Intermediate Economic Statistics. New York,Wiley,1968. p 259-265.

3. Farrar,D.E. & Glauber,R.R.Multi-collinearity in Regression Analysis: the Problem Revisited. *Review of Economics and Statistics*. v.49,1967. p 92- 107.

4. Klein,L.R. Introduction to Econometrics. London,Prentice-Hall, 1963. p 64-101.

5. *ibid*.

6. Theil,H. Specification Errors and the Estimation of the Economic Relationships. *Review of International Statistical Institute*.v.25,1957. p 41-51.

7. Frisch,R.Statistical Confluence Analysis by means of Complete Regression Systems.slo,University Economics Institute, 1934. .

8. Koutsoyiannis,A. Theory of Econometrics. New York,Macmillan,1987, p 239.
