The previous two chapters were devoted to discussion of the object of this research project and a survey of existing literature on research works etc., utilising financial ratios and concerning share price behaviour. In this chapter we shall attempt to devise a concrete design of our research methodology and to mention the data used for statistical analyses and their source, keeping in mind the object described in Chapter I about this research project. For this purpose the present chapter will be divided into three sections, viz.,

1. The Hypothesis and the Model of Relationship
2. The Technique and Tests Used and
3. The Data Used.

These sections are detailed hereafter:

1. The Hypothesis and the Model of Relationship

The hypothesis on which we shall chiefly concentrate our attention in most of our later part of this report is:

The behaviour of investors towards a particular company (as revealed in stock exchange quotations of equity share prices) is dependent, mainly, upon the performance of the company.

The performance itself is broadly composed of two parts, viz.,

(1) the operating efficiency and
(2) the financing effectiveness.

So, symbolically, the relationship of the above mentioned dependence can be delineated strictly in functional form, for a specific company, as

\[ B = f(O,F,S). \]
Here, $B =$ explicit behaviour of the investors in the market towards the company,

$O =$ the operating efficiency,

$F =$ the financing effectiveness and

$S =$ the proxy for some excluded exogeneous variables.

Among the excluded variables the main categories are

(i) macro-economic variables, e.g., money supply, price index as a measure of inflation, the foreign trade and foreign exchange position, the internal as well as external debt position, the stock price index and so on, for the economy of a country as a whole;

(ii) the socio-political variables, e.g., the social value system, war or peace, improved or worsened diplomatic relation with an influential state, the political regime of a democratic state etc. and

(iii) the variables representing the characteristics of other systems, e.g., drought, flood, earthquake, heavy or poor harvests and the like (including the time factor).

For a more extensive formulation some of the variables in categories (ii) and (iii) above may be considered as dummy (with 0 or 1 value) variables.

Our principal concern is to get overall performance score of a company. The said score may also be used to evaluate the performance of the management in charge of the stewardship of the organisation as a whole. Therefore, we are keenly interested in the internal forces $O$ and $F$ of the functional relationship.
depicted in equation (1). Consequently, our revised (noisy) formulation reduces to

\[ B = f(O,F) : S \text{ acting as noise.} \quad (2) \]

The behaviour of investors, in general term, is a qualitative phenomenon and hence this attribute is not measurable quantitatively. But the explicit behaviour of the investors in the market towards the company is spontaneously reflected via the prices of equity shares, more specifically, reflected through

1. average share price for a period,
2. fluctuation of prices in that period indicating risk attached to such prices and
3. change of average price of share over the previous period.

On the other side, the operating efficiency of a company is recognised by different criteria for different operating activity areas, traditionally expressed by some ratios. These criteria are:

(a) turnover of capital,
(b) stock turnover,
(c) operating efficacy of direct expenses,
(d) operating efficacy of indirect expenses and
(e) invested long-term capital mix, i.e., working vs. fixed capital relation.

Similarly, financing effectiveness may be said to be composed of

(i) owners' equity as a multiple of equity share capital,
(ii) financed long-term capital mix, i.e., financial leverage and
(iii) the measure of dividend policy, i.e., the payout ratio.

It should be further mentioned that the current ratio and the yield (dividend-price ratio) could have been taken into consideration. But the current ratio reflects only a very short-term liquidity and the yield has no direct bearing on company level performance (but on investors' personal interest); so, they are excluded.

Eventually, it is quite evident that our model for depicting, in concrete form, the original hypothesis by means of equation (2) converges to the formulation of an objective relation between aspects of equity share prices (in the following period), the dependent variables (1) to (3) above on the one hand and operating efficiency and financing effectiveness (in the current period), the variables representing the criteria (a) to (e) and (i) to (iii) above respectively, on the other. The variables in the dependent set are three, i.e., more than one. Therefore, the suitable model, as per statistical terminology, is canonical correlation analysis and hence the method will be, hereafter, called the canonical correlation analysis for corporate overall performance. The canonical correlation analysis technique which is an extension of the multivariate regression analysis technique, is tersely described with the statistical (sampling) tests to be performed, in the next section. The theory of canonical correlation and tests for the strength of such correlation may be had in Anderson (1958, 1984), Gittins (1985), Srivastava & Yau (1989), Hair Jr. et al. (1990), Pillai & Gupta (1969) and Pearson &
Hartley (1972), for further details and statistical tables.

2. **The Technique and Tests Used**

(a) **The Canonical Correlation Analysis Technique**

As already said at the end of the previous section, the canonical correlation is an extension of the multivariate regression technique with more than one dependent variables.

Let \( y \) and \( x \) be two random vectors with \( p \) and \( q \) elements respectively. Further, \( y \) and \( x \) will be respectively considered as the vectors of dependent and independent variables, such that, \( p < q \). Suppose, there are \( N \) observations (cases), such that, \( N > p + q \). Suppose further that \( a \) and \( b \) are other two vectors (representing the weights) with \( p \) and \( q \) elements respectively and the values of their elements (the weights) are so chosen that the correlation between \( a'y \) and \( b'x \) is maximum. More simply, if the scalars

\[
Y = a'y = \sum_{i=1}^{p} a_i y_i \quad \text{and} \quad P = b'x = \sum_{j=1}^{q} b_j x_j
\]

(here, \( a, b, x \) and \( y \) are the elements of vectors \( a, b, x \) and \( y \) respectively), then, the correlation co-efficient (\( r \), say) of \( Y \) on \( P \) is maximum. The inherent property of the canonical correlation analysis process is so that maximising the correlation co-efficient, \( r \), will automatically find out the weight vectors \( a \) and \( b \).

It may be worth mentioning that \( a \) as well as \( b \) is not a fixed vector, rather, any constant multiple of \( a \), say, \( ka \), and \( kb \),
will also maximise the correlation for same value of r. Therefore, a can be normalised arbitrarily. Generally, a is so normalised that the length of the vector becomes unity. But as in our model the first element of y, \( y_1 \), is the value of average equity share price, the value of the first element of a, \( a_1 \), will be taken as unity. Actual statistical process of calculation is described below:

Suppose, S is the sample sum of squares and cross products matrix of total p+q variables (dependent variables are taken first) and it is partitioned as

\[
S = \begin{pmatrix}
S_{11} & S_{12} \\
S_{21} & S_{22}
\end{pmatrix}
\]  

(5)

here, subscripts 1 and 2 are used to indicate dependent and independent sets of variables respectively. Maximising r is equivalent (using the rule of derivation and maximisation of matrix equations) to solving matrix equation

\[
(S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} - r^2 I)a = 0
\]

(6)

here, I is the identity matrix and o is the null vector of order p. But this matrix equation is a system of homogeneous equations. To get a non-trivial solution the following equation is to be solved first, viz.,

\[
|S_{11}^{-1} S_{12} S_{22}^{-1} S_{21} - r^2 I | = 0
\]

(7)

The left hand side of equation (7) is a determinant. By expansion of this determinant we get a polynomial of degree p which is known as
characteristic polynomial and the resulting equation (7) as characteristic equation. Eventually, by solving this polynomial equation we get p values for the characteristic root (latent root or eigen value) \( r^2 \). It may be noted that all values of \( r^2 \) will be positive and less than (or equal to) unity and hence we get p-values for the correlation co-efficient \( r \). In our model of relationship we shall get three values for the correlation co-efficient, say, \( r_1 \), \( r_2 \) and \( r_3 \), in order of magnitude. But as we are interested with the highest correlation, we shall take the largest value \( r_1 \) for further consideration. Now, if the correlation co-efficient (largest) is statistically significant (by applying any test described in the next sub-section) at some predetermined level, we can conclude that the explanatory set of variables \( x \) is able to account for the variation of the dependent set of variables \( y \) jointly (in spite of the objections against such accounting ratio variables specified in first part of section 1 of chapter II).

After solving for \( a \) from the matrix equation (4) and normalising this vector, as per our policy specified earlier, we get the vector \( b \) from the following relation

\[
b = \frac{1}{r} S_{22}^{-1} S_{21} a
\]

As mentioned earlier \( b \) is our desired weight vector obtained from an objective analysis process. Again,

\[
P = \sum_{j=1}^{q} b_j x_j , \text{ the Eqn. (4)}
\]

is our weighted sum type of overall performance score, an market
based accounting number. It will be, hereafter, called overall performance score, \( P \). It may be noted that variances of \( P \) and \( Y \) are same and the relation between \( Y \) and \( P \) holds as

\[ Y = rP + e, \quad (9) \]

e is the error term, such that, \( \sum e = 0 \).

Following the similar calculation procedure as equation (4) (weighted sum, using those weights of corresponding ratio variables) we shall be able to get operating efficiency score, \( O \), and financing effectiveness score, \( F \). As the relationship is linear in form, obviously

\[ P = O + F. \quad (10) \]

Suppose, there are \( m \) numbers of ratio-variables determining operating efficiency, then

\[ O = \sum_{i=1}^{m} b_i x_i \quad \text{and} \quad (11) \]

\[ F = \sum_{k=m+1}^{q} b_k x_k. \quad (12) \]

For facilitating easy comparison \( x_i \)'s may be measured from their respective arithmetic means.

(b) Statistical Tests to Test the Strength of Canonical Correlation

To test the significance of the canonical correlation coefficient derived from the above statistical process the following parametric tests may be performed, namely,

(1) Likelihood ratio-test by calculating Bartlett's
approximate $\chi^2$ (chi-square) —

$$\chi^2_a (pq) = \frac{s}{N + 1} \log_e \Lambda$$

(13)

here, Wilks' $\Lambda = \prod_{k=1}^{s} (1 - r_k^2)$ ;

(14)

(2) Marriott's (1952) approximate $\chi^2$-test —

$$\chi^2_a (d) = \frac{s}{N + 1} \log_e (1 - r_1^2)$$

(15)

here, $d = p + q - 1 + \frac{s}{3}[(p - 1)(q - 1)]$ and

(16)

(3) Union-intersection test (or gcr-test) due to Roy —

In this test the statistic $r_1^2$ is tested, as it has the greatest characteristic root distribution (gcr), $\theta_a (s, m, n)$. The critical values of this distribution for 0.05 and 0.01 levels of significance are tabulated by Harris (1975, Table A5, p.300).

In all the cases the symbols have the following meanings:

- $N$ = No. of observations or cases,
- $p$ = No. of dependent variables,
- $q$ = No. of independent variables,
- $r_k^2 = k$-th characteristic root, in order of magnitude, of the characteristic equation (7) in p.47 of this report,
- $s$ = minimum of $p$ and $q$, in our case $s = p$,
- $m = \frac{s}{2}(|p - q| - 1)$,
- $n = \frac{s}{2}(N - p - q - 2)$,
- $a$ = level of significance for the (subscripted) test statistic and quantities within brackets with $\chi^2$ etc. represent degrees of freedom for the concerned test statistics.
As we are interested in \( r^2_1 \), we shall use the tests (2) or (3) especially the simpler test (2) above, for testing our original hypothesis mentioned at the first part of section 1 of this chapter. Other tests required for special purpose (e.g., for testing the contribution of a sub-set of variables or a particular variable) will be discussed in the corresponding chapters (e.g., in Chapter X to test the significance of leverage and/or dividend pay-out ratios).

3. The Data Used

Time series data for different years for forty Indian companies and cross-sectional data for 5 years for some of the thirty one companies in General Engineering Industry of India (5 companies are in common to both the categories) are collected from Bombay Stock Exchange Official Directory. All data for share-prices and figures from accounting statements are available from different issues of several volumes of the said directory. Within each volume of the directory only the companies having suitable data are selected for study. Therefore, in fact all companies with required data are selected from the volumes covered (mentioned in Table 1 appended to this chapter). However, the data for macro-economic variables required for section 4 of Chapter X are collected from the different issues of the Reserve Bank of India Bulletin (monthly).

Lists of primary items of data collected from the said directory are appended to this report in Appendix A for companies
in different sectors. For different chapters different definitions of variables (within the periphery of the criteria stated in section 1 of this chapter) are used. Therefore, those definitions will be provided in details in the corresponding chapters, relating them with the items of the primary data. However, the items in Appendix A are defined under 'Explanatory Notes' of Volume 1 of the said directory.

The companies selected for time series study belong to different industries. Their distribution according to industrial classes is given in Table 1 which follows shortly. It may be noted that though State Bank of India is a statutory corporation, it has shares held by the public. So, this corporate body (but not a company under The Companies Act, 1956) is also selected for study.

NOTES: 1. It is a long-run supplement but not a contesting alternative to EMH, rather individual price changes (even due to prospective green or red signals from inside or outside environment) within a period may be explained by EMH.

2. In July, 1993 the Accounting Standards Board (1993) issued a statement on 'Operating and Financial Review' which was designed as a formulation and development of best practice. This statement was intended to have persuasive rather than mandatory force. In this statement the Operating and Financial Review (OFR) was viewed as a framework for the directors to discuss and analyse the business's performance and the factors underlying its results and financial position, in order to assist the users to assess for themselves the future potential of the business.
Table 1

INDUSTRIAL CLASSES OF COMPANIES (OR CORPORATE BODIES HAVING SHARES QUOTED IN STOCK EXCHANGES) SELECTED FOR TIME SERIES CANONICAL CORRELATION STUDIES FOR CORPORATE OVERALL PERFORMANCE

<table>
<thead>
<tr>
<th>Serial</th>
<th>Sector / Industry / Category</th>
<th>No. of Companies</th>
<th>Vol. No. of BSE-Directory</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

A. Manufacturing Companies

1. Under Engineering Industry
   (1) Aluminium
   (2) Iron & Steel
   (3) Electrical
   (4) General

2. Under Cotton Textile Industry
   (5) Cotton Spinning & Weaving Mills

B. Non-Manufacturing Companies

1. Banking (State Bank of India)
2. Trading
3. Investment & Finance
   (1) Proper
   (2) Hybrid Type
4. Transporting : Shipping

TOTAL NUMBER OF COMPANIES ETC. 40