CHAPTER NINE

PREDICTING SICKNESS: A COMPANY LEVEL ANALYSIS

9.1 Introduction

At the company level, the readily available financial data are the financial data published in the balance sheet and profit and loss account of a company. Under the Indian Companies Act, 1956, it is mandatory that a public limited company will publish the balance sheet and profit and loss account at the end of each financial year duly audited by a chartered accountant. It is also mandatory that these balance sheets and profit and loss accounts are to be kept in the public domain. Whether a company is maintaining a proper financial health or whether a company is moribund should be reflected in the financial statements of the company as documented under the balance sheet and profit and loss account. It is true that in many cases, the companies take recourse to window dressing in connivance with unscrupulous chartered accountants so that the financial reality with respect to the concerned company remains underfocused or even undermined. Only the experts with long experience in analysing the balance sheets and profit and loss accounts can understand the reality.

There should be a few crude balance sheet and/or profit and loss account items (or ratios) which would transmit proper signals with respect to the financial health of a company. Even without expertise in the area of the analyses of balance sheets and profit and loss accounts, one should get a feel about the financial health of the companies. Since such

\[102\] In terms of section 210 of the Companies Act, 1956, the Board of Directors of every company is required to place before the Annual General Meeting a balance sheet at the end of each financial year and a profit and loss account for that period. Further, section 215 of the Companies Act, 1956 provides that the balance sheet and the profit and loss account of every company duly approved by the Board of Directors shall be submitted to the auditors for their report.

\[103\] Section 219 of the Companies Act, 1956 provides that every company shall send a copy of the balance sheet and profit and loss account along with auditors’ report to each member at least twenty one days before the Annual General Meeting. Further, in terms of section 159 of the Companies Act, 1956, every company shall file with the Registrar of Companies a return containing particulars which inter alia includes balance sheet and profit and loss account for the period ended as on the last day of the financial year.
items and ratios can be gathered from the information kept in the public domain, even an untrained person should make an assessment and take a decision on whether to invest in stocks and debentures of the said company. Received wisdom in the subject of financial analysis is that no such crude indicator exists. In this dissertation, we have taken this issue for detailed analyses. After considering the relevant literature and identifying the macro ratios of sickness and corresponding micro ratios, we tried to approach this issue with a view to addressing this problem. To be precise, we want to find some crude indicators of sickness from the balance sheets and profit and loss accounts of the companies that might predict sickness—such crude ratios that can be rationalised on the basis of theoretically robust macro and micro ratios of sickness. The issues relating to finding macro and micro ratios have been discussed at length in the previous Chapters of the dissertation. In the pages that follow, we would try to find the corresponding (crude) ratios that can be readily derived, even by a layman from the balance sheet and profit and loss account items of the company. Once we identify the ratios, we expect to develop a model that can be utilised for predicting company level sickness straight from the balance sheet and profit and loss account of a company. This is expected to be user-friendly and can be applied with respect to any Indian company, the financial data of which are available in the public domain.

9.2 Identifying the Relevant Balance Sheet and Profit and Loss Account Ratios

9.2.1 Balance Sheet Items

In order to identify these ratios, we considered first all the items that appear in the liability side of the balance sheet except items constituting shareholders fund or net worth. Understandably, net worth is the barometer that reflects the health of the company. In India, sickness in official parlance is described in terms of net worth of a company. If a company reports negative net worth in the balance sheet at the end of any financial year, technically the company is sick104 and it is mandatory that the company would be referred to the BIFR provided the company is registered for not less than five years. However, this

104 As defined in the Sick Industrial Companies (Special Provisions) Act, 1985, ‘sick industrial company’ means an industrial company (being a company registered for not less than five years) which has at the end of any financial year accumulated losses equal to or exceeding its entire net worth.
does not meet our issue. Our issue is not to discuss the *ex post* situation with respect to sickness. We want to discuss the *ex ante* situation, the situation when the company is not technically sick, that is the net worth of the company is not negative but there are indications that there is a probability of the company getting sick in future. This is important particularly in the context of individual’s decision-making with respect to investing in a company. This is also important in the context of financial management of a company which needs a signal so that it might take some pre-emptive measures before these are referred to the BIFR.

Coming back to the balance sheet items, we started with all the items except the items which constitute shareholders’ fund or net worth. After scrutinising the items, as they appear in the PROWESS database, we realise that one particular item in the database, namely, fixed asset is problematic. We observed that the fixed assets of some of the selected companies were revalued. As one knows, revalued fixed assets do not reflect the true picture of the financial status of a company which is why net worth does not include revaluation reserve in the liability side. Again, gross fixed assets were depreciated which is normal in a running company. But then, the depreciation amount was calculated by various companies in various ways in various years following Accounting Standard Six (AS 6) which allows much flexibility. In fact the true worth of a company often remains undisclosed just because the depreciation sum is manipulated at the wish of the Board of Directors of the company. Since we are to identify the signal for sickness, we found it prudent to exclude fixed assets from the list of explanatory variables. The argument that we furnish is that ‘fixed assets’ as explanatory variable will not be reliable with respect to the Indian companies because of the permitted laxity in accounting norms following Accounting Standard Six\(^{105}\). With respect to the other items in the balance sheet, we did not find any problem and therefore, all of them were taken as they appeared in the balance sheet for finding the suitable predictors for sickness of a company.

\(^{105}\) In terms of revised Accounting Standard (AS) 6, a company is allowed to charge depreciation either by straight line method or by written down value method. The company is also allowed to change the method of calculating the amount of depreciation. However, the method of calculating depreciation and the amount of depreciation charged in each accounting period should be disclosed in the financial statements for that accounting year along with the disclosure of (i) depreciation methods used; and (ii) depreciation rates or the useful lives of the assets, if they are different from the principal rates specified in the statute governing the enterprise.
9.2.2 Profit and Loss Account Items

With respect to the items of profit and loss account of a company, we did not have any problem. We, therefore, decided to include all the items in the list of items from which the proper predictors will be identified.

9.2.3 Normalisation of the Balance Sheet and Profit and Loss Account Items

How to identify the relevant items from the balance sheet and profit and loss account? Since ROIC and CR were found to be the best ratios that can go with segregation of the company in terms of net worth being positive or negative, the items which have strong correlation with these ratios might serve as suitable predictors. The task, therefore, was to run two regressions. In one, where ROIC will be the dependent variable and the other where CR will be the dependent variable. Explanatory variables or independent variables would be the balance sheet items (except net worth and fixed assets) and the items in the profit and loss account.

While organising the data for running the regressions, it was understood that the items in the list of explanatory variables are to be normalised. If we take them straight from the balance sheets and profit and loss accounts, inter-company variations in terms of size of the company—— reflected either in terms of the total assets or total turnover (total income) would affect the results. Normalisation was, therefore, done by taking the ratios of all these items—— first by dividing all the items by total income and secondly by total assets.

Normalisation by total assets may be criticised on the ground that total assets include fixed assets which are not dependable—— at least in the context of the present database. Admittedly, total assets are calculated as gross fixed assets minus depreciation minus value of revaluation plus current assets. Total assets thus would remain contaminated by the undependable entries under depreciation and the value of revaluation. We would, however, submit that total asset is not an explanatory variable in our scheme of analyses. It serves as a denominator with respect to other items in the balance sheet and profit and
loss account. Understatement or overstatement with respect to fixed assets will not then affect any particular explanatory variable. Whatever be the bias, that will be evenly distributed over all the items in the balance sheet and profit and loss account. The relative strength of an explanatory variable as predictor will in no way be affected by this normalisation exercise.

For the empirical exercise, we thus had the data normalised by total income and total assets for all the nineteen variables for ten successive years from 1995. Before running the regressions we checked the quality of the data with respect to regression analysis. It was observed that time series data possessed high degree of autocorrelation as indicated by the relevant value of the test statistic. We, therefore, decided to construct a set of cross section data which would not suffer from autocorrelation. Averages of ten year values of the variables were calculated and these were used in regression and further analyses.

The basic data set, i.e., averages of nineteen ratios used for econometric analyses at the company level study has been provided in Appendix Table 9.1 of this dissertation.

9.3 Regression Analyses

Two sets of regressions were run, one with CR as the dependent variable and the other with ROIC as the dependent variable. The structure of the linear regressions is given as below:

\[ y_1 = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_{19} x_{19} \] \hspace{1cm} (i)
\[ y_2 = b_{02} + b_{12} x_1 + b_{22} x_2 + \ldots + b_{19,2} x_{19} \] \hspace{1cm} (ii)

The test statistic is Durbin-Watson d statistic. Durbin–Watson test is performed to find out serial correlations between errors in regression models. Specifically, it tests whether adjacent residuals are correlated, which is useful in assessing the assumption of independent errors. The test statistic may vary between 0 and 4. A value of 2 means the residuals are uncorrelated. A value less than 2 denotes positive correlation. In our analyses, we got the Durbin–Watson statistic as 1.908 which clearly indicates positive correlation.
where, \( y_1 \) = average CR for ten years
\( y_2 \) = average ROIC for ten years
and,
\( x_1 \) = Raw materials and stores / Total income
\( x_2 \) = Power and fuel / Total income
\( x_3 \) = Salaries and wages / Total income
\( x_4 \) = Raw materials and stores / Total assets
\( x_5 \) = Power and fuel / Total assets
\( x_6 \) = Salaries and wages / Total assets
\( x_7 \) = Short term borrowings / Total assets
\( x_8 \) = Long term borrowings / Total assets
\( x_9 \) = Current liabilities and provisions / Total assets
\( x_{10} \) = Interest / Total income
\( x_{11} \) = Total income / Total assets
\( x_{12} \) = Other fixed costs / Total income
\( x_{13} \) = Other fixed costs / Total assets
\( x_{14} \) = Current assets / Total income
\( x_{15} \) = Current assets / Total assets
\( x_{16} \) = Short term borrowings / Total income
\( x_{17} \) = Long term borrowings / Total income
\( x_{18} \) = Current liabilities and provisions / Total income
\( x_{19} \) = Total borrowings / Total assets
and,
\( b_{ij} \) = \( i \)th regression coefficient with respect to \( j \)th regression equation
\( I = 1, 2, \ldots \ldots 19 \)
\( J = 1, 2 \)

\( b_{01} \) and \( b_{02} \) are the intercept terms.

We have run the regression by stepwise estimation method as our objective is to get
exclusively the combination of those balance sheet and profit and loss account ratios that
might be responsible to explain the maximum variation of each of the dependent
variables, namely, CR and ROIC.
9.3.1 Findings of Regression Analysis with CR as the Dependent Variable

We have first considered CR as the dependent variable. We regressed CR on the nineteen independent variables. We found Model 3 as the best fit model, where three out of nineteen ratios, namely, total borrowings/total assets ($X_{19}$), power and fuel/total income ($X_2$) and current liabilities and provisions/total assets ($X_9$) explained above 71 per cent variation of the CR (Adjusted R square being 0.711). Summarised results are given in Table 9.1. We should point out that expected signs of the $\beta$ coefficients have been realised with respect to $X_2$ and $X_9$. With respect to $X_{19}$, the expected sign has not been realised because of the existence of some sick companies where the profit was negative in spite of very high borrowings compared to their assets. However, variation was high among the realised values of $b_{19,1}$ as the difference between unstandardised and standardised $b_{19,1}$ indicate. We should also report that variation was high with respect to $b_{9,1}$ as well.

Table 9.1: Summarised Results of Regression Analysis with CR as the Dependent Variable

<table>
<thead>
<tr>
<th>Model Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

a Predictors: (Constant), $X_{19}$  
b Predictors: (Constant), $X_{19}$, $X_2$  
c Predictors: (Constant), $X_{19}$, $X_2$, $X_9$
Chapter Nine

Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardised Coefficients</th>
<th>Standardised Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>1.554</td>
<td>0.104</td>
<td>14.93</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-1.812</td>
<td>0.18</td>
<td>-0.712</td>
<td>-10.041</td>
</tr>
<tr>
<td>2</td>
<td>(Constant)</td>
<td>1.744</td>
<td>0.085</td>
<td>20.406</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-1.605</td>
<td>0.144</td>
<td>-0.631</td>
<td>-11.11</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>-4.872</td>
<td>0.623</td>
<td>-0.444</td>
<td>-7.82</td>
</tr>
<tr>
<td>3</td>
<td>(Constant)</td>
<td>1.803</td>
<td>0.089</td>
<td>20.35</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-1.558</td>
<td>0.144</td>
<td>-0.612</td>
<td>-10.833</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>-4.238</td>
<td>0.682</td>
<td>-0.386</td>
<td>-6.211</td>
</tr>
<tr>
<td></td>
<td>X9</td>
<td>-0.412</td>
<td>0.195</td>
<td>-0.132</td>
<td>-2.106</td>
</tr>
</tbody>
</table>

9.3.2 Findings of Regression Analysis with ROIC as the Dependent Variable

Next, we considered ROIC as the dependent variable which was regressed on the nineteen independent variables taken from the balance sheet and profit and loss account items. We found Model 4 as the best fit model where four ratios, namely, total borrowing/total assets (X19), power and fuel/total income (X2), current liabilities and provisions/total assets (X9) and other fixed costs/total assets (Xo) explained above 68 per cent variation of the ROIC. Summarised results as given in Table 9.2 indicate that in the best fit model all beta coefficients, except $b_{19,2}$, had the expected signs in the estimated regression equation. With respect to $b_{19,2}$ the existence of (-) sign and the large difference between the unstandardised and standardised beta is explained in the same way as in case of the previous regression model. We should also report that standard error is found to be high with respect to $X_{13}$ as well.
Table 9.2: Summarised Results of Regression Analysis with ROIC as the Dependent Variable

Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.702</td>
<td>.493</td>
<td>.488</td>
<td>.05392</td>
</tr>
<tr>
<td>2</td>
<td>.798</td>
<td>.636</td>
<td>.629</td>
<td>.04591</td>
</tr>
<tr>
<td>3</td>
<td>.818</td>
<td>.669</td>
<td>.659</td>
<td>.04404</td>
</tr>
<tr>
<td>4</td>
<td>.826</td>
<td>.682</td>
<td>.669</td>
<td>.04337</td>
</tr>
</tbody>
</table>

a Predictors: (Constant), X19
b Predictors: (Constant), X19, X2
c Predictors: (Constant), X19, X2, X9
d Predictors: (Constant), X19, X2, X9, X13

Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardised Coefficients</th>
<th>Standardised Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Beta</td>
<td>Std. Error</td>
<td>Beta</td>
<td>Zero-order</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>.131</td>
<td>.011</td>
<td>11.733</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-.189</td>
<td>.019</td>
<td>-.702</td>
<td>-9.768</td>
</tr>
<tr>
<td>2</td>
<td>(Constant)</td>
<td>.148</td>
<td>.010</td>
<td>14.966</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-.170</td>
<td>.017</td>
<td>-.632</td>
<td>-10.148</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>-.447</td>
<td>.072</td>
<td>-.385</td>
<td>-6.178</td>
</tr>
<tr>
<td>3</td>
<td>(Constant)</td>
<td>.158</td>
<td>.010</td>
<td>15.766</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-.162</td>
<td>.016</td>
<td>-.603</td>
<td>-9.966</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>-.342</td>
<td>.077</td>
<td>-.295</td>
<td>-4.425</td>
</tr>
<tr>
<td></td>
<td>X9</td>
<td>-6.794E-02</td>
<td>.022</td>
<td>-.206</td>
<td>-3.071</td>
</tr>
<tr>
<td>4</td>
<td>(Constant)</td>
<td>.144</td>
<td>.012</td>
<td>11.860</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>X19</td>
<td>-.150</td>
<td>.017</td>
<td>-.557</td>
<td>-8.713</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>-.317</td>
<td>.077</td>
<td>-.273</td>
<td>-4.110</td>
</tr>
<tr>
<td></td>
<td>X9</td>
<td>-7.269E-02</td>
<td>.022</td>
<td>-.220</td>
<td>-3.316</td>
</tr>
<tr>
<td></td>
<td>X13</td>
<td>.172</td>
<td>.086</td>
<td>.126</td>
<td>1.995</td>
</tr>
</tbody>
</table>
9.3.3 Conclusion from the Regression Analyses

What should we conclude at the end of the regression exercises with respect to CR and ROIC? Given the data set, the stepwise regression which was run with the help of SPSS package was to suggest the best linear regression that could be estimated with respect to the data. What we observed was that the best linear regressions as found from the data set were explaining about 71 per cent and 67 per cent respectively of the behaviour of the explained variables. The explanatory variables associated with these two lines of best fit might, therefore, be taken as the variables which would explain the behaviour of ROIC and CR to a reasonably large extent. One can, therefore, rely on these explanatory variables for CR and ROIC based modelling on company level sickness. These variables as we have already observed are as follows: total borrowings/total assets ($X_{19}$), power and fuel/total income ($X_2$), current liabilities and provisions/total assets ($X_9$) and other fixed costs/total assets ($X_{13}$).

Out of the nineteen balance sheets and profit and loss account items, therefore, only four items would matter in predicting sickness of a company. If the balance sheet and profit and loss account items are reliable, a practitioner would identify only four items reporting the status of the company, namely, the status of the company with respect to dependence on borrowings, status as regards efficiency in power and fuel consumption and the situation with respect to current liabilities and provisions that describes the commitment of the company in terms of liquidity. The regression analysis points out that other fixed costs in the nature of regular commitments of the company also matter in determining the financial health of the company.

The summary of the balance sheet and profit and loss account items is available in the form of net worth of a company – the single most important indicator of the financial health of a company which is widely used for assessing the financial situation of a company. For predicting sickness, however, the net worth based modelling will not be appropriate. One cannot predict whether the net worth of a company will be worse or whether it will improve in the next year just on the basis of the net worth figures of the
previous years. This is so because net worth is the end product of various balance sheet and profit and loss account items. The historical information on net worth of a company shall have to be disaggregated for developing any model that might predict whether the net worth in the future would be worse or better. If the individual balance sheet and profit and loss account items are to be analysed for predicting sickness, the modelling would involve considering not net worth but individual items. This rationalises our approach of considering individual items on the basis of certain ratios and finding out the most sensitive items in the balance sheet and profit and loss account. On the basis of the panel data, we have observed that there are four items as mentioned above, which are most sensitive to the financial health of a company included in the panel.

We are also convinced that the statistical exercise based on the panel data could not produce any alternative set of balance sheet and profit and loss account items that might be used for predicting sickness. This is so because the regressions had been run stepwise and the models for the best fit were automatically generated. The extent of robustness (captured by R-square and the values of standard errors of estimates) that could be realised, had been the best possible that one might find, given the panel data. We are aware of the fact that the expected sign of $\beta$ coefficient was not realised in case of one variable ($X_{19}$). This was due to the volatility with high absolute values with respect to ‘total borrowings to total assets’ in the data set that we were working with. However, this is what is expected with respect to the balance sheets of certain companies that are not maintaining financial discipline (which is why they are prone to sickness). This fact of life has to be taken into consideration while modelling on the basis of the historical data set. We would keep this point in mind while utilising these items for developing model for predicting sickness. Before we take up the issue of predicting sickness by utilising the most sensitive four ratios as mentioned above, we propose to find out a discriminant score or $Z$ score with these four ratios as independent variables. The $Z$ score thus developed would classify the companies into healthy or sick groups.
9.4 Classification of Companies by Z score

In this section, we shall first classify the companies by Z score to be calculated on the basis of the four selected balance sheet and profit and loss account ratios. The initial data for calculating the Z score and finding the cut-off would be done on the basis of panel data with respect to one hundred companies—the data set with which we are working. Classification by applying this cut-off will also be discussed. We could then utilise this cut-off score of Z (and also the estimated values of the parameters) for validation of the hypothesis that this methodology can predict whether the composition of the group will remain undistorted. The issue of predicting sickness on the basis of the existing data with respect to a particular company is a different proposition. This issue we shall take up in the next section by means of binary logistic regression.

9.4.1 Method of Classification

The method of classification was derived from multiple discriminant analysis (MDA) with a discriminant function\(^{107}\) in the form \(Z = a_1v_1 + a_2v_2 + a_3v_3 + \ldots + a_nv_n\).

The discriminant function transforms the values of various variables to a single discriminant score or Z value, which is then used to classify the object, where

- \(a_1, a_2, \ldots, a_n\) are discriminant coefficients and
- \(v_1, v_2, \ldots, v_n\) are independent variables.

The MDA computes the discriminant coefficients, \(a_j\), where the independent variables \(v_j\) are the actual values and \(j = 1, 2, 3, \ldots, n\).

The technique of MDA had been used by the researchers in the past as an effective tool for prediction of corporate sickness. Prof. E.I. Altman is considered to be the pioneer in this area of research\(^{108}\). In all the models, including the model developed by Altman to

---

\(^{107}\) Discriminant function is a linear combination of the independent variables used in the classification of group membership. The estimated value of the discriminant function is the discriminant score (Z score). The Z score is calculated for each object and used in conjunction with the cut-off score to identify the group to which an object belongs.

classify the companies into healthy and sick groups, various balance sheet and profit and loss account items (ratios) were taken as explanatory variables. What is then the difference in our approach and the approach adopted by Altman? We shall discuss this issue first before we discuss about the findings of our discriminant analyses.

9.4.2 Altman's Classification Method and Our Approach

In order to determine the parameters of the system, Altman considered the data of a select set of companies for five successive years and on the basis of the average performance in terms of explanatory variables, a cut-off level was identified and the classification of the companies was done accordingly. These explanatory variables (various financial ratios) were, however, selected on the basis of some a priori considerations from the balance sheets and profit and loss accounts of selected companies and finally a cut-off score was found out which was used to classify the companies into healthy and sick groups. In our research, we have taken a different approach. We have not chosen the financial ratios arbitrarily. We have first considered certain macro economic ratios. We carried out various statistical analyses, namely, rank analysis, scatter plot, convergence and divergence analysis, cluster analysis and tested whether these macro ratios can capture the heterogeneity or homogeneity in performance of the industry (NIC two digit level) groups. We found that two macro ratios, namely, CR and ROIC can segregate the ‘good performing’ industry groups from ‘bad performing’ industry groups. We then found out the corresponding CR and ROIC from the micro or company level data. We found that these corresponding micro ratios, namely, ROIC and CR had power to differentiate the ‘good performing’ (healthy) companies from the ‘bad performing’ (sick) companies—healthy or sick defined in terms of ‘net worth’ criterion. As we wanted to develop a model which can be used even by a non-expert, we identified balance sheet and profit and loss account items (ratios) that can be found out easily from the balance sheet and profit and loss account of a company. Considering all the items of balance sheet and profit and loss account of a company, we got nineteen ratios that can be considered as explanatory
variables. As a next step, we found out only those ratios out of these nineteen ratios which can significantly explain the variation of ROIC and CR which were found to be the best ratios for segregating the 'good performing' companies from the 'bad performing' companies. By running regressions, we found four financial ratios, namely, total borrowings/total assets ($X_{19}$), current liabilities and provisions/total assets ($X_9$), power and fuel/total income ($X_2$) and other fixed cost/total assets ($X_{13}$) could explain significantly the variation of CR and ROIC. We thus considered only these four financial ratios as independent variables in multiple discriminant analysis. The model that we propose to develop for prediction of financial health of a company would thus be free from the problems of arbitrariness.

What we would highlight is the fact that Altman's Z score does not assign a probability value to a particular company getting sick in future. It provides the necessary information on the probability of the calculated Z score being able to classify a group of companies into healthy or sick groups. The discussion on sickness widely utilises Altman's methodology. Admittedly, the methodology is robust. One should not, however, ignore the fact that the Altman's Z score does not in any way help us understand the future course of a particular company. On the basis of historical data set, one can, however, point out whether the classification is reliable, i.e., whether the number of companies belonging to a particular group is expected to remain in the same group in future. If the alteration in number is not high, one can reasonably expect a particular company to remain in the same group in future. Altman's methodology is not designed to assign a probability value of a company getting sick or otherwise.

9.5 Findings of Simultaneous Discriminant Analysis

We have done a two-group (healthy and sick) simultaneous discriminant analysis of the companies to predict the status of a company – whether healthy or sick, considering four
balance sheet and profit and loss account ratios, namely, total borrowings/total assets (X19), current liabilities and provisions/total assets (X9), power and fuel/total income (X2) and other fixed costs/total assets (X13) as discriminating variables. The programme was run with the help of SPSS package. The data set on the basis of which discriminant analysis was performed is given in Appendix Table 9.2. Among alternatives, first one canonical discriminant functions were used in the analyses. Binary correlation coefficients among the explanatory variables being low, the selection of the explanatory variables appears to be satisfactory. The multivariate measures of overall model fit of the discriminant function which became significant displayed a canonical correlation of 0.848. This implies that above 71 per cent of the variation of the dependant variable has been accounted for by this model. From structure matrix, we obtained the relative contribution of each of the independent variables to the discriminant function, i.e., the variance that an independent variable shares with the discriminant function. Table 9.3 shows the outcome of the discriminant analysis. We observe that amongst the four ratios, total borrowing / total assets (X19) has more discriminating power followed by power and fuel/total income (X2) and other fixed costs/total assets (X13). Current liabilities and provisions/total assets (X9) has least discriminating power.
Table 9.3: Outcome of Discriminant Analysis

### Eigenvalues

<table>
<thead>
<tr>
<th>Function</th>
<th>Eigenvalue</th>
<th>% of Variance</th>
<th>Cumulative %</th>
<th>Canonical Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.562</td>
<td>100.0</td>
<td>100.0</td>
<td>.848</td>
</tr>
</tbody>
</table>

### Standardised Canonical Discriminant Function Coefficients

<table>
<thead>
<tr>
<th>X_2 (power and fuel/total income)</th>
<th>.728</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_9 (current liabilities and provisions/total assets)</td>
<td>.084</td>
</tr>
<tr>
<td>X_13 (other fixed cost/total assets)</td>
<td>(-).305</td>
</tr>
<tr>
<td>X_19 (total borrowings/total assets)</td>
<td>.900</td>
</tr>
</tbody>
</table>

### Structure Matrix

<table>
<thead>
<tr>
<th>Function</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_19</td>
<td>0.655</td>
</tr>
<tr>
<td>X_2</td>
<td>0.402</td>
</tr>
<tr>
<td>X_13</td>
<td>-0.32</td>
</tr>
<tr>
<td>X_9</td>
<td>0.238</td>
</tr>
</tbody>
</table>

### Binary Correlation Coefficients Among the Explanatory Variables

<table>
<thead>
<tr>
<th></th>
<th>Power and fuel/total income</th>
<th>Current liabilities and provisions/total assets</th>
<th>Other fixed costs/total assets</th>
<th>Total borrowings/total assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power and fuel/total income</td>
<td>1</td>
<td>0.338</td>
<td>0.075</td>
<td>-0.368</td>
</tr>
<tr>
<td>Current liabilities and provisions/total assets</td>
<td>0.338</td>
<td>1</td>
<td>0.134</td>
<td>-0.056</td>
</tr>
<tr>
<td>Other fixed costs/total assets</td>
<td>0.075</td>
<td>0.134</td>
<td>1</td>
<td>-0.091</td>
</tr>
<tr>
<td>Total borrowings/total assets</td>
<td>-0.368</td>
<td>-0.056</td>
<td>-0.091</td>
<td>1</td>
</tr>
</tbody>
</table>
The discriminant function that we obtained from the above model is:

\[ Z = 0.728 X_2 + 0.084 X_9 - 0.305 X_{13} + 0.900 X_{19} \]

where, \( X_2 \) = power and fuel/total income,
\( X_9 \) = current liabilities and provisions/total assets,
\( X_{13} \) = other fixed costs/total assets,
\( X_{19} \) = total borrowings/total assets

Once we obtain the discriminant function and values of the discriminant coefficients of the explanatory variables, our next step is to determine the ‘cut-off’ score (Z score). The company whose Z score will be above the ‘cut-off’ score will be classified as a sick company and the company whose Z score will be below the ‘cut-off’ score will be identified as a healthy company. We have done this exercise in the next section.

### 9.6 Determination of the ‘Cut-off’ Score

As a first step, we calculated Z scores for each of the hundred companies selected in our panel data. The output is shown in Appendix Table 9.3. In calculating the ‘cut-off’ score, we arranged the hundred Z scores in ascending order. Then we calculated the ‘cut-off’ score as the average of the Z score of fiftieth company (\( Z_{50} \)) and fifty first company (\( Z_{51} \)) arranged in ascending order. Thus, we got the ‘cut-off’ score as 0.432342. Calculation for finding out the ‘cut-off’ score is given in Appendix Table 9.4.

With the help of the computed cut-off score, i.e., 0.432342, we have classified each of the one hundred companies selected in our panel data into either healthy group or sick group on the basis of the Z score of an individual company. The company whose Z score is above the ‘cut-off’ score, i.e., 0.432342 would be considered as a bad performing (sick) company and the company whose Z score is below the ‘cut-off’ score would be considered as a good performing (healthy) company. Our objective is to find out the robustness of the discriminant model that we have developed and shown in the earlier section. In the next section we have done this exercise by constructing a classification matrix.
9.7 Construction of Classification Matrix

We used a simple format for constructing the classification matrix. The format is framed as follows:

<table>
<thead>
<tr>
<th>Predetermined grouping (a priori grouping)</th>
<th>Grouping classified according to the discriminant score (Z score)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
</tr>
<tr>
<td>Healthy</td>
<td>C_h</td>
</tr>
<tr>
<td>Sick</td>
<td>W_s</td>
</tr>
</tbody>
</table>

C stands for correct classification and W stands for misclassification. The sum of C_h and C_s equals the total correct classification and when this sum is divided by the total number of companies classified (one hundred in the case of our sample for empirical analysis), gives the measure of correctness of the MDA in classifying companies according to their a priori groupings. This percentage measures the per cent of variation of the dependent variable explained by the independent variables.

For constructing the classification matrix, we calculated the Z score for each of the fifty companies which are pre-classified as healthy companies by applying the discriminant function. Similarly, Z score for each of the fifty companies which are pre-classified as sick companies were calculated. On the basis of these calculated Z scores, each company
from *a priori* grouping was given a status of either a healthy or a sick company depending on the individual Z score of that particular company. Finally, the classification matrix was constructed.

The classification matrix for the hundred companies selected in our panel data for empirical analysis is as follows:

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
<td>Sick</td>
</tr>
<tr>
<td>Healthy</td>
<td>46</td>
<td>4</td>
</tr>
<tr>
<td>Sick</td>
<td>4</td>
<td>46</td>
</tr>
</tbody>
</table>

From the classification matrix, we find that the discriminant model developed by us could classify 92 per cent of the total sample correctly. Error in both the groups is only 8 per cent. The results, therefore, are encouraging.

**9.8 Validation of the Model**

As one knows, quantitative exercise on modelling requires the validation of the proposed model with respect to another set of data to judge the acceptability of the proposed model. We checked the validity of the model developed by us by considering a panel data of another one hundred companies from the PROWESS database, comprising of fifty good performing (healthy) companies and fifty bad performing (sick) companies. The
data set was prepared on the basis of same net worth criterion. The panel data of fifty healthy companies was taken from the good performing industry groups at NIC two digit level, namely, IC 22, IC 26 and IC 30. Net worth of each of these companies was positive in 1995 and showed an increasing trend over the ten years time period from 1995 to 2004. On the other hand, the panel data of fifty sick companies was taken from the bad performing industry groups at NIC two digit level, namely, IC 23, IC 24, IC 28, IC 31, IC 33, IC 35–36 and IC 37. Net worth of these companies was positive in 1995 and became negative at least in the last year and showed a declining trend over the ten years period from 1995 to 2004. List of holdout sample companies is given in Appendix Table 9.5 and Appendix Table 9.6.

For validation of the model which we have developed, we have performed the analysis in two stages. In the first stage, we have checked the classification accuracy of the companies within the hold out sample. For doing this exercise, we first calculated the Z scores for each of the hundred observations of the holdout sample by applying the discriminant function developed by us, i.e.,

\[ Z = 0.728 X_2 + 0.084 X_9 - 0.305 X_{13} + 0.900 X_{19} \]

where,

- \( X_2 = \) power and fuel/total income,
- \( X_9 = \) current liabilities and provisions/total assets,
- \( X_{13} = \) other fixed costs/total assets,
- \( X_{19} = \) total borrowings/total assets

We then compared the Z score with the cutting score of 0.432342 and assigned the status, whether healthy or sick, to each of the hundred companies. The companies whose Z score was above 0.432342 were marked as sick and those with Z score of less than 0.432342 were noted as healthy. We did this for each of the five years from 2000 to 2004. The classification matrix thus found out is given in Table 9.4. Z scores computed for each of the holdout sample companies from 2000 to 2004 are given in Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11.
### Table 9.4: The Classification Matrix for Holdout Sample for the Years 2000-2004

<table>
<thead>
<tr>
<th>Year</th>
<th>Predetermined group</th>
<th>No. of companies</th>
<th>Classified according to Z score</th>
<th>Classification accuracy (percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Healthy</td>
<td>Sick</td>
</tr>
<tr>
<td>2000a</td>
<td>Healthy</td>
<td>50</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Sick</td>
<td>50</td>
<td>9</td>
<td>41</td>
</tr>
<tr>
<td>2001b</td>
<td>Healthy</td>
<td>50</td>
<td>44</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Sick</td>
<td>50</td>
<td>4</td>
<td>46</td>
</tr>
<tr>
<td>2002c</td>
<td>Healthy</td>
<td>50</td>
<td>43</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Sick</td>
<td>50</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>2003d</td>
<td>Healthy</td>
<td>50</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Sick</td>
<td>50</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>2004e</td>
<td>Healthy</td>
<td>50</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Sick</td>
<td>50</td>
<td>1</td>
<td>49</td>
</tr>
</tbody>
</table>

Source: Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11

Notes:

a. In 2000, out of five good performing companies which are classified as bad performing by the cutting score, one company falls into shadow region and two, out of nine bad performing companies which are classified as good performing by the cutting score, fall into shadow region.

b. In 2001, all the four bad performing companies which are classified as good performing by the cutting score, fall into shadow region.

c. In 2002, out of seven good performing companies which are classified as bad performing by the cutting score, one falls into shadow region.

d. In 2003, out of five good performing companies which are classified as bad performing by the cutting score, one falls into shadow region.

e. In 2004, out of five good performing companies which are classified as bad performing by the cutting score, one falls into shadow region.

We thus observed that classification accuracy varied from about 86 percentage to about 95 percentage. One would observe from Table 9.4 that the level of accuracy was more during the terminal year than that during the earlier years. As a next step, we have seen how many companies that were classified as either healthy or sick in 2000 on the basis of our Z score have maintained the same status in 2004, i.e., after a lapse of four years. Same exercise was performed with reference to base year 2001, 2002 and 2003. Results are shown in Table 9.5, Table 9.6, Table 9.7 and Table 9.8.
Table 9.5: Predictive Power of the Model (Base Year=2000)

<table>
<thead>
<tr>
<th>Status</th>
<th>Percentage of correct prediction</th>
<th>No. of good performing companies becoming bad performing</th>
<th>No. of bad performing companies becoming good performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before four years (2000 &amp; 2004)</td>
<td>86</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Before three years (2000 &amp; 2003)</td>
<td>89</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Before two years (2000 &amp; 2002)</td>
<td>87</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Before one year (2000 &amp; 2001)</td>
<td>92</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11

Table 9.6: Predictive Power of the Model (Base Year=2001)

<table>
<thead>
<tr>
<th>Status</th>
<th>Percentage of correct prediction</th>
<th>No. of good performing companies becoming bad performing</th>
<th>No. of bad performing companies becoming good performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before three years (2001 &amp; 2004)</td>
<td>92</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Before two years (2001 &amp; 2003)</td>
<td>93</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Before one year (2001 &amp; 2002)</td>
<td>95</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

Source: Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11

Table 9.7: Predictive Power of the Model (Base Year=2002)

<table>
<thead>
<tr>
<th>Status</th>
<th>Percentage of correct prediction</th>
<th>No. of good performing companies becoming bad performing</th>
<th>No. of bad performing companies becoming good performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before two years (2002 &amp; 2004)</td>
<td>97</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Before one year (2002 &amp; 2003)</td>
<td>98</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11
Table 9.8: Predictive Power of the Model (Base Year=2003)

<table>
<thead>
<tr>
<th>Status</th>
<th>Percentage of correct prediction</th>
<th>No. of good performing companies becoming bad performing</th>
<th>No. of bad performing companies becoming good performing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before one year (2003 &amp; 2004)</td>
<td>97</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: Appendix Tables 9.7, 9.8, 9.9, 9.10 and 9.11

From the results, one may observe that predictive power of our model is satisfactory. One should, however, point out that the predictive power of our model has been tested on the basis of another historical data set. The implication is that a post facto scenario has been verified by the model. To what extent a particular company has the probability of getting sick (or otherwise) in future, i.e., during a period for which the historical data does not exist, cannot be ascertained by this model. This, we submit, is the basic limitation of the analysis of sickness on the basis of Altman Z score. Such exercises do discriminate between the healthy group and a sick group of industries. Whether a particular company will belong to the predetermined group (healthy or sick) in future cannot be ascertained unless the ex post scenario is available.

The practitioners, however, might not be interested in the academic exercise of Z score based classification. For the practitioner the issue is “what is the probability of a particular company becoming sick in future?” If the probability is low, the practitioner would adopt a bullish strategy in the stock market. In case the probability of getting sick is high, the practitioner would adopt a bearish strategy in the stock market. Since the classification by Z score in a particular year does not, in any way, make a probability based statement on the future health of a particular company, such classification exercises would in effect provide no signal as regards its future financial health. In order to address such an issue, we propose to use another statistical tool—a tool that would help us address the issue directly. With this tool we should be able to find the probability of a particular company getting sick in future on the basis of historical data set. This tool is a variation of the tool of regression analysis. Known as logistic regression, this tool would be utilised in developing a predictive model on the basis of the panel data that we are working with.
9.9 Predicting Sickness: Binary Logistic Regression

To recapitulate, our objective is to find out the probability of a company becoming healthy or sick on the basis of the four balance sheet and profit and loss account ratios that have been found to have discriminating power. These ratios are total borrowings/total assets ($X_{19}$), current liabilities and provisions/total assets ($X_9$), power and fuel/total income ($X_2$) and other fixed cost/total assets ($X_{13}$). The statistical tool that we have used for predicting sickness is binary logistic regression. Logistic regression is a statistical tool for modelling the relationship between a dependent (response) variable and a set of independent (explanatory) variables when the dependent variable is categorical. In binary logistic, the categorical variable takes only two values, namely, 1 and 0. Logistic regression is specifically designed to predict the probability of an observation being in either of the two groups. For example, suppose that the dependent variable $Y$ takes the values 1 and 0, and one wants to model the probability of $Y=1$ as a function of some explanatory (independent) variables. The logistic regression would approach this problem by considering the ‘odd’ ratio (probability of $Y=1$ divided by the probability of $Y=0$) as the response variable and then constructing a linear relation between the log of odds and a set of explanatory variables, i.e., the variables that might determine the ‘odd’ against $Y=1$.

The technical literature on logistic regression is available elsewhere. As the literature suggests, the logit (logistic regression) model is based on cumulative logistic probability function. If $Y_i = \alpha + \beta X_{i}$, then cumulative logistic probability function will be specified as:

\[ P_1 = \frac{1}{1 + e^{-(\alpha + \beta X_i)}} \]

---

which ultimately gives a form:

\[ e^{y_i} = \frac{P_i}{1 - P_i} \]

Taking natural log on both side,

\[ Y_i = \ln \left( \frac{P_i}{1 - P_i} \right) \]

i.e., \( \alpha + \beta X_i = \ln \left( \frac{P_i}{1 - P_i} \right) \)

If \( \frac{P_i}{1 - P_i} \) is taken as the dependent variable, it will be simply logarithm of odds that a particular event will take place.

In our case, this particular event is getting healthy. In terms of a logit model, the probability of odds against getting healthy is expressed as a linear function of the independent variables \((X_i)_s\). In our case, the model takes four variables as the independent variables. These variables are four balance sheet and profit and loss account ratios expressed in quantitative terms (not as binary variables, i.e., variables expressed in ‘0’ and ‘1’). The rationale for applying the logit model is that it transforms the problem of predicting probabilities within a \((0,1)\) interval to the problem of predicting the odds of events occurring within the range of entire set of real values.

In this model, there are no indicator variables. All the variables are quantitative variables expressed in terms of four chosen ratios. Being healthy is the ‘event’ and the probability of ‘event’, i.e, the probability of being healthy is expressed as

\[ \text{Probability (event)} = \frac{1}{1 + e^{-z}} \]
where, $Z$ is the linear combination and is equal to $\beta_0 + \beta_1 X_{19} + \beta_2 X_2 + \beta_3 X_9 + \beta_4 X_{13}$

where, 
- $X_{19} = \frac{\text{Total borrowings/total assets}}{}$
- $X_2 = \frac{\text{Power and fuel/total income}}{}$
- $X_9 = \frac{\text{Current liabilities and provisions/total assets}}{}$
- $X_{13} = \frac{\text{Other fixed cost/total assets}}{}$

The number of independent variables in this case is four. The probability of ‘event’ not occurring is estimated as:

$$\text{Probability (no event)} = 1 - \text{Probability (event)}$$

The odd ratio i.e., the odd of being healthy is expressed as:

$$\text{Odd} = \frac{\text{Probability (event)}}{\text{Probability (non event)}}$$

This odd ratio for a variable tells us the change in odds for a case when the value of that variable increases by one.

From the given data set with respect to one hundred Indian companies for which the balance sheet and profit and loss account data were available, we planned to run the binary logistic regression.

As one knows, the estimation exercise with respect to the logistic regression does not follow the same procedure, as in case of ordinary least square estimation. Estimation exercise in such a case where ordinary least square technique cannot be used follows a maximum likelihood estimation procedure. Given the data set, the software packages can be utilised for finding the maximum likelihood estimators with respect to $\beta_0, \beta_1, \beta_2, \beta_3$ and $\beta_4$. The statistical test for examining the robustness of the model would be different from what we ordinarily do with respect to least square estimation. First, let us consider the issue of reliability of the estimated $\beta$s. Since the maximum likelihood estimators are known to be asymptotically normal, the analog of the regression $t$ test can be applied, because, in this case, the ratio of the estimated coefficients to their estimated standard errors follows a normal distribution. If we wish to test the significance of all or a subset
of the coefficients in the logit model when maximum likelihood is used, a test using the 
chi-square distribution replaces the usual F test. To be precise, a likelihood ratio ($\lambda$) is 
defined as $\lambda = L_0/L_{\text{max}}$

where, $L_0$ = initial value of the likelihood function 
and $L_{\text{max}}$ = maximum of the same function

The appropriate test follows directly from the fact that

$$-2 \log \lambda = -2 \log \left( \frac{L_0}{L_{\text{max}}} \right)$$

follows a chi-square distribution of $k$ degrees of freedom where,

$k$ = no. of parameters in the equation (other than the constant term).

Ordinarily, the software packages provide Wald Chi-square values with respect to 
estimated $\beta$s for testing the reliability of the estimated $\beta$s.

Next is the issue of goodness of fit. The goodness of fit in case of ordinary least square is 
measured by the value of $R^2$ (Adjusted). In this case, no such direct measure of $R^2$ is 
possible. Various measures of goodness of fit analogous to $R^2$ have been suggested. In 
our case, we used SAS package where Cox and Snell $R^2$ and Nagelkerke $R^2$ (Max-
rescaled R-Square) have been used for testing the goodness of fit. The package also 
provides the Hosmer and Lemeshow goodness of fit. The results of the logit run with 
respect to SAS are given in Appendix Table 9.12. We should point out that the model 
was run with respect to the same database on which the discriminant analysis was done 
(see Appendix Table 9.2).

9.9.1 Results of Binary Logistic Regression Analysis

First, we shall report the findings on the test of goodness of fit with respect to the 
empirical model. The Cox and Snell R-Square value (0.6988) and the Max-rescaled
Chapter Nine

R-Square value (0.9317) signify that our model is likely to be a good fit one. The result of Hosmer and Lemeshow goodness of fit test indicates a strong association between the predicted probabilities and the observed responses. This is further supported by high values of ‘Per cent Concordant’ (99.3 per cent), ‘Somers’ D’ (0.987), ‘Gamma’ (0.989) and a low value of ‘Per cent discordant’ (0.6 per cent) along with a low value of ‘Per cent Tied’ (0.2). We observe from the result of testing global null hypothesis: \( \beta = 0 \) that the chi-square value of the likelihood ratio statistic is significant even at 1 per cent level of significance. This implies that the four ratios, namely, total borrowings/total assets \( (X_{19}) \), current liabilities and provisions/total assets \( (X_9) \), power and fuel/total income \( (X_2) \) and other fixed cost/total assets \( (X_{13}) \) simultaneously can identify the status of a company, whether healthy or sick with more than 99 per cent probability of being true. Analysing the classification table, we observe that with probability of 0.72, 97 per cent of the selected companies are correctly classified into their respective predetermined group by our model. We observe that out of fifty predetermined healthy companies, probability of only one company (H2-4) remaining as healthy is less than 1 per cent. On the other side, out of fifty predetermined sick companies, probability of only six companies, namely, S1-1, S1-3, S2-1, S2-9, S3-3 and S3-7 losing the status of remaining as sick company varies from 29 per cent to 54 per cent.

With respect to the estimated values of the \( \beta \) coefficients, we observe that the Wald Chi-square values of the estimators are satisfactory. Thus, with respect to \( \beta_0 \), the Wald Chi-square value is 8.5069 and the probability that the calculated value will differ from 8.5069 in any other experiment is only 0.0035. For the other estimated \( \beta_k \) i.e., \( \beta_1, \beta_2, \beta_3 \) and \( \beta_4 \), the estimated values are likely to be true with more than 95 per cent probability.

9.9.2 Signs of the Estimated \( \beta \) Coefficients

The binary logistic regression model, applied over our data set generates the following estimated probability of a company being healthy.

\[
P (\text{Healthy}) = \frac{1}{1 + e^{-Z}}
\]

where, \( Z = 9.5909 - 14.3800 X_{19} - 46.9134 X_2 - 15.5178 X_9 + 126 X_{13} \)
As the results of the empirical exercise indicate, the expected signs of the estimated $\beta$s have been realised. For example, the sign of $\beta_1$ is negative. The implication is that higher the value of total borrowings/total assets ($X_{19}$), lower the odds that the event 'healthy' will occur, i.e., lower is the probability that a company will become healthy. Again, sign of $\beta_2$ is negative which implies that higher is the value of power and fuel/total income ($X_2$), lower is the odds that the event 'healthy' will occur, i.e., lower is the probability that a company will become healthy. Sign of $\beta_3$ is also negative. This implies that more is the value of current liabilities and provisions/total assets ($X_9$), less is the odds that the event 'healthy' will occur, i.e., lower is the probability that a company will become healthy. Finally, sign of $\beta_4$ is positive which implies that higher is the value of other fixed cost/total assets ($X_{13}$), higher is the odds that the event 'healthy' will occur, i.e., higher is the probability that a company will become healthy. These are consistent with the expectation of a financial analyst. Suppose, the value of the ratio of total borrowings to total assets ($X_{19}$) for a company is different for two financial years. For the year $Y_{t+1}$, $X_{19}$ is higher than $X_{19}$ for the year $Y_t$. There could be three situations under which this may happen. First, value of total assets for both the years would remain same, but the value of total borrowings for the year $Y_{t+1}$ would be more than the value of total borrowings for the year $Y_t$. In this case, increase in the amount of total borrowings would result in increase in the amount of the interest expenses which would ultimately reduce profit after tax for the year $Y_{t+1}$. Consequently, the rate of return on investment [i.e., $\frac{(PAT + \text{Interest})}{\text{Total Assets}}$] would also decline in the year $Y_{t+1}$. Second, rate of increase in the amount of total borrowings could be more than rate of increase in the amount of total assets during the year $Y_{t+1}$. This would result in decline in the rate of return on investment in the year $Y_{t+1}$ from the rate of return on investment for the year $Y_t$. Lastly, the amount of total borrowings could remain same, but the value of total assets for the year $Y_{t+1}$ would be less than the value of total assets for the year $Y_t$. This would happen when there is erosion in net worth of the company during the year $Y_{t+1}$. In any of these three situations, there exists a possibility that the company's performance during the year $Y_{t+1}$ would be bad as compared to the previous year. If the trend in increase of $X_{19}$ continues in the subsequent years, the possibility of the company getting sick would be more.
Similarly, if value of the ratio of current liabilities and provisions to total assets \( (X_9) \) of a company is more for the year \( Y_{t+1} \) compared to the value of \( X_9 \) for the year \( Y_t \), it implies that the liquidity of the company has reduced during the year \( Y_{t+1} \). Less liquidity of a company would result in shortage of working capital. For shortage of working capital, there is a possibility that the company would not be in a position to utilise its capacity fully. Further, for shortage of working capital, the company would not be in a position to hold stock of raw materials and stock of finished goods at optimum level. All these factors would result in less turnover, higher operating cost and less profit margin during the year \( Y_{t+1} \) compared to the previous year \( Y_t \). If this trend continues in the future years, probability of the company getting sick would be more.

Cost of power and fuel is a variable cost. Higher value of the ratio of power and fuel to total income \( (X_2) \) indicates higher operating cost compared to total income. If the value of \( X_2 \) is more during the year \( Y_{t+1} \) compared to the value of \( X_2 \) in the previous year \( Y_t \), it implies that the company is deploying inefficient technology or the plant and machinery are getting obsolete. Thus, the value of 'contribution' which is equal to the excess of value of total income over the value of total variable cost would be less during the year \( Y_{t+1} \) compared to the previous year \( Y_t \). Less would be the amount of contribution; less would be the ability of the company to service fixed overheads and interest expenses. Thus, if the value of \( X_2 \) of a company is more during the year \( Y_{t+1} \) compared to the value of \( X_2 \) during the year \( Y_t \), it indicates that the performance of the company has deteriorated. If this trend continues in the future years, the probability of the company getting sick is more.

Other fixed costs considered for calculating the value of \( X_{13} \) (other fixed costs / total assets) are mainly comprised of cost for marketing, advertisement and distribution. These costs are incurred mainly for sales promotion. Only healthy companies can incur such costs. If the amount of such costs incurred by a company during the year \( Y_{t+1} \) is more than the amount incurred during the year \( Y_t \), it is expected that the value of sales would be more during the year \( Y_{t+1} \) compared to the previous year \( Y_t \). This would mean more capacity utilisation in the year \( Y_{t+1} \) compared to the previous year \( Y_t \). With increase in
capacity utilisation the cost of other fixed overheads does not normally increase. In such a situation, profit margin of the company increases. If this trend continues in the future years, probability of the company getting sick becomes much less.

9.9.3. Predicting the Financial Health of a Company

On the basis of the empirical results with respect to one hundred companies we thus have a statistically sound model that predicts probability of a company being 'healthy'. The model suggests -

\[
\text{Probability of the event (being healthy)} = \frac{1}{1 + e^{-Z}}
\]

with \( Z = 9.5909 - 14.3800 X_{19} - 46.9134 X_2 - 15.5178 X_9 + 126 X_{13} \)

This model serves as a predictor with respect to any company belonging to the universe from which the panel data have been prepared. In the context of the companies for which the balance sheet and profit and loss account data are available, one may work out, without much effort, the four relevant ratios, namely, total borrowings /total assets (\( X_{19} \)), power and fuel/total income (\( X_2 \)), current liabilities and provisions/total assets (\( X_9 \)) and other fixed costs/total assets (\( X_{13} \)). These may be placed in the model and the corresponding \( Z \) value may be calculated. With the \( Z \) value, one can immediately calculate the probability of a company being healthy. Let us take an example. We select a company, namely, A (real name changed) from the PROWESS database of CMIE and find out the relevant four ratios, namely, total borrowings /total assets (\( X_{19} \)), power and fuel/total income (\( X_2 \)), current liabilities and provisions/total assets (\( X_9 \)) and other fixed costs/total assets (\( X_{13} \)) on the basis of its balance sheet and profit and loss account. Value of these ratios is as follows:

\[
\begin{align*}
X_{19} &= 0.00124 \\
X_2 &= 0.00699 \\
X_9 &= 0.24373 \\
X_{13} &= 0.00187
\end{align*}
\]
Thus,
\[ Z = 9.5909 - (14.3800 \times 0.00124) - (46.9134 \times 0.00699) - (15.5178 \times 0.24373) + (126 \times 0.00187) \]
i.e., \[ Z = 5.698868276 \]
Therefore,
\[ P = \frac{1}{1 + e^{-5.698868276}} = 0.99 \]
Thus, the probability of the company A being healthy is above 99 per cent.

Let us consider another company, namely, B (real name changed) from the PROWESS database of CMIE and find out the relevant four ratios, namely, total borrowings /total assets (\(X_{19}\)), power and fuel/total income (\(X_2\)), current liabilities and provisions/total assets (\(X_9\)) and other fixed costs/total assets (\(X_{13}\)) on the basis of its balance sheet and profit and loss account. Values of these ratios are as follows:

\[
\begin{align*}
X_{19} &= 0.61700 \\
X_2 &= 0.09880 \\
X_9 &= 0.22334 \\
X_{13} &= 0.01800
\end{align*}
\]
Thus,
\[ Z = 9.5909 - (14.3800 \times 0.61700) - (46.9134 \times 0.09880) - (15.5178 \times 0.22334) + (126 \times 0.01800) \]
i.e., \[ Z = -5.114349372 \]
Therefore,
\[ P = \frac{1}{1 + e^{-5.114349372}} = 0.006 \]
Thus, the probability of the company B being healthy is below 1 per cent, i.e., the probability of the company B being sick is above 99 per cent.

Thus, we conclude that the probability of remaining healthy or otherwise is sensitive to the four ratios that we have identified. The model suggested by us thus appears to be satisfactory. From the balance sheet and profit and loss account, one may find very easily the numerical values of the four suggested ratios and applying these to the model that we have developed in the context of the Indian scenario, one may find the probability that the company would remain healthy or would become sick.

The odd ratio that gives the probability of event as a ratio to probability of non event may also be derived from the same model. For example, odd ratio with respect to the case where total borrowings /total assets \( X_{19} = 0.00124 \), power and fuel/total income \( X_{2} = 0.00699 \), current liabilities and provisions/total assets \( X_{9} = 0.24373 \) and other fixed costs/total assets \( X_{13} = 0.00187 \) is

\[
\frac{\text{Probability of event}}{\text{Probability of non event}} = 99
\]

As a numerical value, this indicates that the odds in favour of being sick are rather low for this company. A change in odd ratio can, however, be calculated thus in the previous example. Let us consider a hypothetical scenario where the values of power and fuel/total income \( X_{2} \), current liabilities and provisions/total assets \( X_{9} \) and other fixed costs/total assets \( X_{13} \) remain unchanged and the value of total borrowings /total assets \( X_{19} \) increases to say 0.20. The resultant P is 0.94 and odd ratio is 15.67. Comparing this with the previous odd ratio, we understand that the other things remaining the same, if total borrowings /total assets \( X_{19} \) increases by almost 161 times, the odd ratio changes to 15.67. The ratio of odds of being healthy when total borrowings /total assets \( X_{19} \) is 0.20 to the same odds when total borrowings /total assets \( X_{19} \) was 0.00124 is about 84 percentage. This ratio for a variable tells us change in the odds for a case when the value of that variable increases by almost 161 times. Needless to say, such an exercise can be
performed by considering a change by 1 per cent, 2 per cent, etc. Typically, this exercise helps one perform a sensitivity analysis with respect to the probability of remaining healthy or getting sick.

**Chapter Nine**

9.10 *Summary and Conclusion*

One would thus observe that the outcomes of our analyses are satisfactory. This signifies that the four balance sheet and profit and loss account ratios, namely, total borrowings /total assets \(X_{19}\), power and fuel/total income \(X_2\), current liabilities and provisions/total assets \(X_9\) and other fixed costs/total assets \(X_{13}\), are good indicators not only for identifying the status of a company – whether healthy or sick, but also for predicting a company being healthy. One has to find out these four balance sheets and profit and loss account ratios for a company and find out the probability of the company being healthy from our predictive model. These ratios can be found out easily from the balance sheet and profit and loss account of a company even by a person having not much knowledge of analyses of financial statements of a company. The predictive model constructed by us would, therefore, be a ‘user-friendly’ model. The model is expected to be helpful for the investors, bankers, lenders and credit rating agencies.
Chapter Nine

References

- Maddala, G S, *Introduction to Econometrics*, John Willey & Sons, Ltd.