IIMB Management Review (2013) 25, 104—114

available at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/iimb

Industrial sickness in India — An empirical analysis

Dilip Kumar Datta [a,b,*]

[a] Sayantan Consultants Pvt. Ltd, Kolkata, India
[b] Indian Institute of Social Welfare and Business Management (IISWBM), University of Calcutta, India

KEYWORDS
Convergence and divergence analysis;
Rank analysis;
Scatter plot analysis;
Cluster analysis;
Centroids;
Adjusted R square;
Discriminant score;
Canonical correlation

Abstract In addressing the core issue of the probability of a company becoming sick in the future and developing predictive models of financial health, this article first classifies Indian industries into two groups - good performing and bad performing - on the basis of ASI data. This classification is based on broadly accepted economic indicators which are translated into corresponding financial ratios. Balance sheet data of 100 companies, evenly drawn from the two groups, is then analysed and the future financial health of these companies is predicted on the basis of a logit model.

Introduction

The *sine qua non* of industrial sickness is that production fails to maintain cost-effectiveness, so much so that the firm fails to meet the sunk cost. Firms that cannot maintain competitive efficiency face the reality of getting sick. It is thus necessary for firms to get a signal that there is a threat to the existing business so that they take up strategies for a turnaround. There is a vast literature on this subject, developed mostly by practising financial experts. However, as we discuss in this paper, existing empirical exercises in industrial sickness following Altman's (1968) seminal work on the subject often fail to predict accurately. This failure could be attributed to an arbitrary selection of ratios for analysis which lack an adequate macro foundation.

As is generally accepted, the industrial climate in which a firm operates is best reflected in macro economic indicators. Firm level analysis on industrial sickness is based on balance sheet data. In order to have the appropriate tool for prediction one has to identify the correspondence between macro and micro indicators of industrial performance. The existing literature on firm level sickness does not recognise this fundamental point adequately. The analyses are done mostly on the basis of arbitrarily selected financial ratios derived from the balance sheet and profit and loss account statements. The element of arbitrariness often tells upon the robustness of the ratios and accordingly, the predictive models fail to predict the future scenario accurately. The present paper aims at addressing these problems by taking a different approach to firm level sickness. We first derive certain parameters from a few broad macro economic indicators generally used by

* Chatterjee International Centre, Suite No. A-5, 9th Floor, 33A, Jawaharlal Nehru Road, Kolkata 700 071, India. Tel.: +91 9831006577, +91 033 2226 6737; fax: +91 033 2226 1846.
E-mail address: sayantan.consultants@gmail.com
Peer-review under responsibility of Indian Institute of Management Bangalore

Production and hosting by Elsevier

0970-3896
http://dx.doi.org/10.1016/j.iimb.2013.02.003

economists for macro level analyses of industrial performance. We use these macro indicators for empirical analysis of the Annual Survey of Industries (ASI) time series (macro level) data on organised manufacturing industries in India. We then develop corresponding financial ratios which are generally used by financial analysts for company level analyses. We would submit that these micro indicators developed by us have been mapped with the macro indicators and can be culled easily from the balance sheet and profit and loss account of a company (at a micro level). These indicators are then used for empirical analyses at company (micro) level with a view to developing discriminant and predictive models.

The paper is planned as follows. A brief review of the literature on empirical models of industrial sickness and their shortcomings are presented in the first section. In the second section we discuss the database and methodology adopted for selecting major industry groups in the Indian manufacturing sector. In the third section we identify the macro indicators of industrial performance (good performing and bad performing) and with this we try to develop the corresponding micro indicators, the indicators that might be used for analysing the balance sheet data of a company (i.e., for a micro unit). We thus find a set of accounting ratios for analysing sickness—the ratios which have a robust macro foundation. This exercise has been carried out in the final part of the paper. The reliability of such indicators has also been tested in this part of the paper. Lastly, we present a summary of the main findings of the paper.

Indicators of sickness: a review of literature

How would financial analysts judge the health of a company? Analysts and policy makers generally use gross financial ratios, which are mostly ad hoc in nature, to assess the health of a company. For example, in India, 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[1] would be referred to the Board for Industrial and Financial Reconstruction (BIFR) as a sick industrial company. The academic exercises on indicators of sickness are largely based on a few financial ratios, selected according to the researcher's perception of company level sickness. It has not been proved that the BIFR approved criteria of sickness are a universally accepted indicator of sickness. Thus L. C. Gupta (1993) used 56 ratios and sought to determine the best set of ratios to predict failure. There are, however, two issues which often remain intermixed in the literature. The first is the issue of finding an indicator of sickness on the basis of performance of a company as reflected in the historical data, mainly the balance sheet and profit and loss accounts data. This at best can provide information on the scenario as it existed. The BIFR considers the problem of sickness from this perspective. The other issue is that of predicting sickness, which is a different exercise. It factors in the probability of getting sick. Even if a company is a good performing one and has not been referred to BIFR, there might exist a probability of it getting sick—a possibility which is not within the purview of discussion of the other set of literature. Models on prediction were developed as a separate set of literature, most important of which is Altman's (1968) model. Altman used multiple discriminant analysis (MDA) to develop a discriminant model based on certain accounting ratios. The ratios were chosen by him on the basis of their popularity in the literature and their relevance to the study. From an initial list of 22 variables, Altman selected five ratios and obtained a discriminant function. Altman's model based on MDA appears to suffer from some limitations arising from the statistical assumptions made. These are: choosing the variables on the basis of their predetermined importance, reducing the number of variables that do not significantly contribute to the overall discriminating model and the selection of *a priori* probabilities. Besides, Altman's Z score does not assign a probability value to a particular company getting sick in future. It only provides information on the probability of the calculated Z score being able to classify a group of companies as healthy or sick on the basis of a historical data set. Whether the companies belonging to a particular group are expected to remain in the same group in future cannot be ascertained from Altman's model. The model is useful only for predictive purposes if the underlying relationship and parameters are stable over time. Otherwise it is only valid for the sample period and it cannot be extrapolated into a subsequent period with the same expected performance. If the alteration in numbers is not high, one can reasonably expect a particular company to remain in the same group in the future. The implication is that a *post facto* scenario has been verified by the model. However, to what extent a particular company has the probability of getting sick (or otherwise) in the future, i.e., during a period for which the historical data does not exist, cannot be ascertained by this model. Such exercises discriminate between the healthy and the 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. Thus the discriminating power of Altman's Z score relates to the characteristics of a particular sample and not to any rationale regarding the actual importance of a set of characteristics in general. 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 with regard to its future financial health. Considering the shortcomings of the existing models we have taken a different approach to predicting industrial sickness in this paper.

Data base and selection of major industries

In order to develop a predictive model using financial ratios which have a macro foundation, industry level and company level data are required. For industry level data, we relied on

[1] Net worth is the sum total of the paid-up capital and free reserves which means all reserves credited out of the profits and share premium account but does not include reserves credited out of re-evaluation of assets, write back of depreciation provisions and amalgamation.

the published data of the Government of India under the Annual Survey of Industries (ASI) and for the company level data, we referred to the PROWESS database compiled by the Centre for Monitoring Indian Economy (CMIE). We selected major industry groups for our empirical analysis with respect to three parameters, namely, value of output, number of workers, and invested capital of respective industry groups. After initial screening, we found that out of 27 industry groups, 15 industry groups accounted for 92.27% of value of output, 90.28% of number of workers and 91.91% of invested capital. Thus, excluding the remaining 12 industry groups (which accounted for only 7.73% of value of total output, 9.72% of number of workers and 8.09% of invested capital) would not affect the result of our analyses of the performance of Indian industries. We thus finally selected 15 such major industry groups for performing our analyses.[2]

Proposed macro indicators of industrial performance: definition, concept, and derivation from ASI data

Productivity, considered as a basic measure of performance of an industry group, often does not reveal an industry's performance in terms of repaying capacity, return on invested capital, cash generation from operation, and management efficiency in controlling its assets and liabilities. This is because improved productivity reflected in terms of reduced cost of production might not mean high return on invested capital as the rate of return depends on the rate of profit which is not necessarily accelerated with a reduction in the cost of production. For example, we find in industry group IC 26 (manufacture of textile products including apparel) capital (physical) productivity is low (1.74) and ranks 7th, but it ranks 2nd with respect to return on invested capital which is high (0.36) among the 15 selected industries. Productivity *per se* thus does not reveal inner strength or weakness of an industry group; the issue is addressed better in terms of a select set of financial parameters. Given this background, we selected seven ratios as indicators for assessing performance of an industry group with respect to its profitability, liquidity, debt-servicing capacity, leverage, and working capital management efficiency. They are: return on invested capital (ROIC), operating cash flow to invested capital (OCF/IC), interest coverage ratio (ICR), debt service coverage ratio (DSCR), leverage ratio (LR) and working capital management efficiency ratio (WCMER) and average of these ratios, namely, composite ratio (CR). These ratios are normally used by financial analysts to assess performance of a company. We derived these ratios from certain macro indicators given in the ASI data, namely, profit, interest, invested capital, fixed capital, working capital, and outstanding loan, and others. In order to maintain conformity and parity, we kept in mind the definition given by financial analysts for these ratios and chos synonymous data from the ASI data in a manner that would maintain definitional parity. For example, we matched "invested capital", an ASI-given item with "total assets", a balance sheet item. Net worth' neither features in the ASI data, nor is it used by economists as a tool for measuring performance of an industry sector at the macro level. However, we consider "net worth" as almost equivalent to fixed capital + working capital − outstanding loan in respect of industry level data.[3]

"Fixed capital", an ASI-given item was considered synonymous with "net block" (gross fixed assets − accumulated depreciation + capital work-in-progress), a balance sheet item. Similarly, "working capital" was considered equivalent to "net working capital" or "net current assets" (current assets − current liabilities) which appears in the balance sheet. In this manner, we got macro level data corresponding to micro level information. We then constructed ROIC which is an indicator of profitability as (profit + interest)/invested capital. Similarly, we took OCF/IC, an indicator of liquidity to be the same as "operating cash flow to total assets" and derived it as (profit + interest + depreciation)/invested capital. Interest coverage ratio (ICR), an indicator for interest servicing capacity was derived as (profit + interest)/interest. Debt service coverage ratio (DSCR), an indicator for debt servicing capacity was derived as (profit + interest + depreciation)/(interest paid + 20% of outstanding loan).[4] Leverage ratio (LR), which is essentially equity-debt ratio and which indicates how much a firm depends on outside borrowings, was derived from the ASI data as (fixed capital + working capital − outstanding loan)/outstanding loan. Working capital management efficiency ratio (WCMER), an indicator for efficiency in regard to management of assets and liabilities, was derived as a ratio between working capital and invested capital. Our process of matching macro and micro

[2] Selected 15 major industry groups are: IC 20—21(manufacture of food products), IC 22 (manufacture of beverages, tobacco, and related products), IC 23 (Manufacture of cotton textiles), IC 24 (manufacture of wool, silk, and man-made fibre textiles), IC 25 (manufacture of jute and other vegetable fibre textiles (except cotton), IC 26 (manufacture of textile products (including wearing apparel), IC 28 (manufacture of paper and paper products and printing, publishing, and allied industries), IC 30 (manufacture of basic chemicals and chemical products (except products of petroleum and coal), IC 31 (manufacture of rubber, plastic, petroleum, and coal products; processing of nuclear fuels), IC 32 (manufacture of non-metallic mineral products), IC 33 (basic metal and alloys industries), IC 34 (manufacture of metal products and parts, except machinery and equipment), IC 35—36 (manufacture of machinery and equipment other than transport equipment), IC 37 (manufacture of transport equipment and parts) and IC 40 (electricity).

[3] In the language of a financial analyst, total assets of a company comprise of net block (gross fixed assets minus accumulated depreciation) investments, and current assets. Liabilities of a company consist of net worth or share holders fund, term loan and current liabilities. Again, according to the accounting equation, assets are equal to liabilities. Thus, net worth of a company is equal to total assets minus total of term loan and current liabilities. Keeping this interpretation of the financial analysts and definition given in the ASI for various items, we find that the sum total of fixed capital and working capital of an industry sector is equivalent to total assets minus current liabilities. If we deduct outstanding loan from this figure, what we get is essentially net worth of an industry sector or an individual industry.

[4] We take 20% of outstanding loan in the denominator as outstanding loan that is normally repaid within a period of five years.

data relied on some basic concepts followed by economists in analysing economic behaviour of an industry; at the same time it did not in any way distort the conceptual framework of a financial analyst. All such variables excepting employee cost have been deflated by the wholesale price index (WPI) with 1982 as the base year. We have used the Consumer Price Index (CPI) to deflate employee cost. There are, however, certain limitations while directly using WPI as deflator. While the ASI classification is based on activities, WPI is based on nature of commodities. We submit that identifying the nature of the commodity grouped under the ASI activity based classification is difficult, if not impossible. At best, one can approximate commodities based on the nature of economic activity which prompted us to use WPI only (except for employee compensation). With respect to company level data, each of these items was normalised first by total assets and then by total income.

Analysis of performance of manufacturing industries in India with respect to seven ratios: the findings

Performance of a firm largely depends on the industry group to which it belongs. It was thus important to test whether the selected seven macro indicators could classify Indian industries into two groups—good performing and bad performing. Once this was tested, they could be utilised for constructing a robust indicator of sickness at the micro (firm) level. We first performed convergence and divergence analysis[5] to find out whether the select ratios could show existence of homogeneity or heterogeneity in performance of the 15 major industry groups. As indicated by the results in terms of the seven ratios[6] (Table 1), these macro indicators show that major industry groups are heterogeneous with respect to their performance. With the objective of finding out which of the 15 industry groups were performing well and which were not, we carried out rank analysis, scatter plot analysis and cluster analysis (K-means method) using SPSS package. Results (Tables 2—4) showed that out of 15 industry groups, values of seven indicators for industry groups IC 22, IC 26, and IC 30 were above all industry average[7] and they had better and more consistent performance. On the other hand, performance of the remaining 12 industries, particularly IC 23, IC 24, and IC 25 with respect to the seven indicators were below all industry average. They have inconsistent (more volatile) behaviour. Based on the seven ratios we classified the 15 industry groups into two groups, namely, good performing and bad performing. IC 22, IC 26 and IC 30 were "good performing" industries and the other 12 industry groups, particularly IC 23, IC 24, and IC 25 were "bad performing" industries. A comparison with BIFR data[8] reveals that several companies from the "bad performing" group had been referred to BIFR as sick industrial companies. Whether the seven ratios could be utilised for constructing a robust indicator of sickness at the micro, i.e., company level had to be tested on the basis of the company level data. We take up such an exercise in the next section.

Table 1 Summary results of test of σ and β convergence among representative industries according to selected ratios.

Parameter	Beta	Coeff. value	Sigma	Coeff. value
C R	D	+0.0022	D	+0.6547
ROIC	D	+0.1065	D	+1.156
OCF/IC	D	+0.0034	D	+1.19
ICR	D	+0.0047	D	+1.92
DSCR	D	+0.0018	D	+1.75
L R	C	−0.0185	C	−5.09
WCMER	C	−0.0482	D	+2.31

Notes: D indicates divergence; C indicates convergence. Parameters: CR: composite ratio; ROIC: return on invested capital; OCF/IC: operating cash flow to invested capital; ICR: interest coverage ratio; DSCR: debt service coverage ratio; LR: leverage ratio; WCMER: working capital management efficiency ratio.

Finding signals of sickness and predicting sickness: a company level analysis

Predicting company level sickness is usually performed in terms of certain ratios derived from balance sheet and profit and loss account data. In this paper, we have selected ratios that correspond and are mapped with the macro indicators which have been tested for merit to classify industries into two groups. Predicting sickness through an arbitrary choice of financial ratios has thus been avoided. The exercise was performed as under:

First, we translated seven macro indicators into micro indicators using balance sheet and profit and loss account items at company level retaining the concept of individual ratio as explained above in the section "Proposed macro indicators of industrial performance: definition, concept and derivation from ASI Data". These corresponding ratios were used for building models to predict sickness at the company level. We collected company level data of 683 registered companies documented by the CMIE in the PROWESS database for a ten-year period (1995—2004). In order to select a set of 100 companies from 683 companies,

[5] The concept of σ *convergence* focuses attention on the dispersion of value of the parameter in question over a cross-section of some comparable units (in our case the 15 Indian industries at two digit NIC level) over a period of time. The units are said to satisfy the condition of σ *convergence* if this dispersion decreases over time. However, the reverse pattern would make one conclude the existence of σ *divergence* among the comparable units. The concept of β convergence relates directly to the growth rates of selected parameters of some group of entities. In β convergence the relationship is negative (slope is negative). In case the slope is positive, it would indicate a case of β divergence.

[6] β *convergence* was tested with the benchmark of the average value of a variable over a period of first five years.

[7] In case of "good performing" industries, ROIC is much above the safety level of 20%; DSCR is twice the loan repayment obligation; ICR is thrice the interest obligation; LR is almost one; WCMER is 0.40; average value of centroids in cluster analysis (1.52) is much above the average value (0.78).

[8] BIFR website: www.bifr.nic.in.

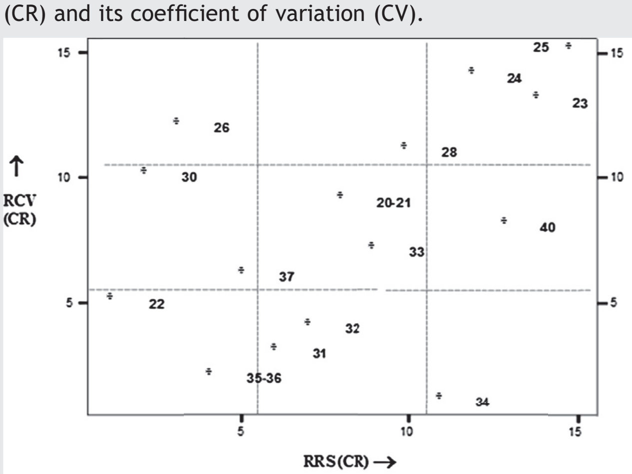
Table 2 Ranks as per selected financial ratios of representative industries in India.

Industry code	ROIC		OCF/IC		ICR		LR		DSCR		WCMER		Composite rank	CR value	
	Rank of mean of ratio	Rank of CV	Rank of mean of ratio	Rank of CV	Rank of mean of ratio	Rank of CV	Rank of mean of ratio	Rank of CV	Rank of mean of ratio	Rank of CV	Rank of mean of ratio	Rank of CV		Rank of mean of ratio	Rank of CV
20—21	7	5	8	8	8	7	11	10	8	6	9	10	9	9	9
22	1	3	1	3	1	4	3	5	1	10	5	9	1	1	5
23	12	13	11	11	13	13	14	15	12	9	13	13	13	14	13
24	14	15	14	14	14	15	4	14	14	14	3	15	12	13	14
25	15	14	15	15	15	14	15	11	15	15	15	14	15	15	15
26	2	9	2	12	3	12	5	12	2	13	2	5	3	3	12
28	10	8	10	7	9	9	8	7	9	7	10	3	10	10	11
30	4	10	4	13	2	10	1	8	3	11	1	12	2	2	10
31	6	2	6	1	6	3	10	2	6	2	8	7	6	6	3
32	8	6	7	4	7	5	7	1	7	3	11	2	8	7	4
33	11	12	12	4	10	11	2	4	11	8	12	6	11	8	7
34	3	4	3	6	11	2	12	9	10	1	6	4	7	11	1
35—36	5	1	5	2	4	1	9	13	4	4	4	1	4	4	2
37	9	11	9	9	5	8	6	3	5	5	7	8	5	5	6
40	13	7	13	10	12	6	13	6	13	12	14	11	14	12	8

Notes: CR: composite ratio; ROIC: return on invested capital; OCF/IC: operating cash flow to invested capital; ICR: interest coverage ratio; DSCR: debt service coverage ratio; LR: leverage ratio; WCMER: working capital management efficiency ratio; CV: coefficient of variation

we defined healthy or sick companies on the basis of a universally accepted criterion, i.e., net worth of a company. On the basis of the reasonably expectation that "good performing" industry groups would have more healthy companies and "bad performing" industry groups would have more sick companies, we scanned the data set from the "good performing" industry groups and selected 50 companies whose net worth was positive both in the initial year and in the terminal year with significant growth rate (β) of net worth (in the form of net worth equal to $\alpha + \beta t$). Similarly, from the "bad performing" group we selected 50 companies whose net worth in the initial year was positive and eventually became negative in the terminal year, and β was significantly negative. Thus, we had a panel of 50 "healthy" companies and 50 "sick" companies. In order to test the robustness of seven micro ratios corresponding to the same macro ratios, we performed cluster analysis with the K-means method as we did for macro level analysis.

Table 3 Scatter plot of ranks in terms of composite ratio (CR) and its coefficient of variation (CV).



Notes:
RRS (CR): Rank as per average rank score in terms of composite ratio
RCV (CR): Rank as per CV of composite ratio

Notes: RRS (CR): rank as per average rank score in terms of composite ratio; RCV (CR): rank as per CV of composite ratio.

Results of the cluster analysis showed that both CR and ROIC had satisfactory differentiating power (low distortion in initial grouping and error percentage of 10).[9] Operating cash flow to invested capital and ICR also recorded a lower percentage error. As these were derived from ROIC, we did not take them into consideration. Results of empirical analyses with ASI data being almost similar, one can conclude that micro ratios are consistent with the macro ratios. Since our main objective was to develop a model for predicting sickness at company level, we needed to identify the relevant ratios from the balance sheet and profit and loss account of a company. Since ROIC and CR are the best ratios that can go with segregation of the company in terms of net worth being positive or negative, the balance sheet ratios that have strong correlation with ROIC and CR might serve as suitable predictors. The task, therefore, is to perform regression analyses.

Regression analyses

We ran two regression analyses by stepwise estimation method—one with ROIC and the other with CR as dependent variable and balance sheet ratios as explanatory variables. We took all the components of balance sheet and profit and loss account of a company and normalised them by dividing first by total income and then by total assets.

[9] With respect to CR, nine "healthy" companies are classified as "sick" companies and one "sick" company is classified as "healthy"; with respect to ROIC, ten "sick" companies are classified as "healthy".

Table 4 Value of centroids of cluster 1, cluster 2 and fifteen major industries.

	Cluster 1	Cluster 2	All the fifteen major industries
ROIC	0.13	0.32	0.17
OCF	0.19	0.38	0.23
ICR	1.47	4.34	2.04
LR	0.61	1.27	0.74
DSCR	0.94	2.34	1.22
WCMER	0.24	0.45	0.28
Average value of centroids	**0.60**	**1.52**	**0.78**

Notes: CR: composite ratio; ROIC: return on invested capital; OCF/IC: operating cash flow to invested capital; ICR: interest coverage ratio; DSCR: debt service coverage ratio; LR: leverage ratio; WCMER: working capital management efficiency ratio.

We thus obtained 19 ratios as explanatory variables.[10] Summarised results (Tables 5 and 6) show that four explanatory variables, 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 associated with lines of best fit explaining about 71% and 68% respectively of the behaviour of dependent variables (CR and ROIC). These accounting ratios out of a total of 19 ratios would matter in predicting sickness of a company. We then proceeded to find out a discriminant score or Z score with these four ratios as independent variables. The Z score thus developed should classify the companies into healthy or sick groups.

Classification of companies by *Z* score

We derive a method of classification from multiple discriminant analysis (MDA) with a discriminant function[11] in the form $Z = a_1 v_1 + a_2 v_2 + a_3 v_3 \ldots\ldots\ldots\ldots a_n v_n$. The discriminant function transforms values of various variables to a single discriminant score or Z value, which is then used to classify the object, where

$a_1, a_2, \ldots\ldots\ldots\ldots, a_n$ are discriminant coefficients and

$v_1, v_2, \ldots\ldots\ldots\ldots, v_n$ are independent variables.

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

We perform a two-group (healthy and sick) simultaneous discriminant analysis of 100 companies which constitute the panel data considering four balance sheet and profit and loss account ratios as stated above as discriminating variables. We use SPSS package for this analysis. Table 7 shows the outcome of the discriminant analysis. The multivariate measures of overall model fit of the discriminant function which became significant display a canonical correlation of 0.848. This implies that above 71% of the variation of the dependant variable is accounted for by this model. Binary correlation coefficients among the explanatory variables is low. It justifies selection of explanatory variables. Amongst four ratios, total borrowing/total assets (X_{19}) has more discriminating power followed by power and fuel/total income (X_2), and other fixed costs/total assets (X_{13}). Current liabilities and provisions/total assets (X_9) has least discriminating power.

The discriminant function obtained from the above model is:

$$Z = 0.728X_2 + 0.084X_9 - 0.305X_{13} + 0.900X_{19}$$

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

In calculating the cut-off score, we calculate Z scores for each of the hundred companies in our panel data and arrange the hundred Z_i scores in ascending order. The cut-off score is the average of the Z score of the fiftieth company (Z_{50}) and

Table 5 Summarised results of regression analysis with CR as the dependent variable.
Model summary.

Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.712	0.507	0.502	0.50267
2	0.835	0.698	0.691	0.39569
3	0.843	0.711	0.702	0.38886

Model 1. Predictor: (Constant), X_{19}.
Model 2. Predictors: (Constant), X_{19}, X_2.
Model 3. Predictors: (Constant), X_{19}, X_2, X_9.

Table 6 Summarised results of regression analysis with ROIC as the dependent variable.
Model summary.

Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.702	0.493	0.488	0.05392
2	0.798	0.636	0.629	0.04591
3	0.818	0.669	0.659	0.04404
4	0.826	0.682	0.669	0.04337

Model 1. Predictor: (Constant), X_{19}
Model 2. Predictors: (Constant), X_{19}, X_2.
Model 3. Predictors: (Constant), X_{19}, X_2, X_9.
Model 4. Predictors: (Constant), X_{19}, X_2, X_9, X_{13}.

[10] 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 and x_{19} = Total borrowings/total assets.

[11] 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.

Table 7 Outcome of discriminant analysis.

Eigenvalues				
Function	Eigenvalue	% Of variance	Cumulative %	Canonical correlation
1	2.562	100.0	100.0	0.848
Standardised canonical discriminant function coefficients				
X_2 (power and fuel/total income)		0.728		
X_9 (current liabilities and provisions/ total assets)		0.084		
X_{13} (other fixed cost/total assets)		(−)0.305		
X_{19} (total borrowings/total assets)		0.900		

the fifty-first company (Z_{51}). In this way, we get the cut-off score as 0.432342. For our model, we construct a classification matrix in a format framed as follows.

Predetermined grouping (*a priori* grouping)	Grouping classified according to the discriminant score (*Z* score)		
	Healthy	Sick	Total
Healthy	C_h	W_h	$(C_h + W_h)$ = total number in the *a priori* grouping (healthy)
Sick	W_s	C_s	$(C_s + W_s)$ = total number in the *a priori* grouping (sick)

C stands for correct classification and W stands for misclassification. The sum of C_h and C_s equals the total correct classification. When this sum is divided by total number of companies classified, we get the measure of correctness according to their *a priori* groupings. Following the classification matrix, we find that the discriminant model developed by us could classify 92% of the total sample correctly. Error in both the groups is only 8%. Results being encouraging, we extend our empirical analysis further to address the core issue of this paper, namely, "What is the probability of a particular company becoming sick in the future?" We perform this exercise by binary logistic regression.

Predicting sickness: 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 and 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 (Hanushek, 1977; Pindyck and Rubinfeld, 1985) 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:

$P_i = \dfrac{1}{1 + e^{-(\alpha+\beta_i)}}$ which ultimately gives a form: $e^{Y_i} = \dfrac{P_i}{1 - P_i}$.

Taking natural log on both side, $Y_i = \ln\left[\dfrac{P_i}{1 - P_i}\right]$

i.e., $\alpha + \beta X_i = \ln\left[\dfrac{P_i}{1 - P_i}\right]$. If $\left[\dfrac{P_i}{1 - P_i}\right]$ is taken as the dependent variable, it will simply be a 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 (Xi_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 an (0,1) interval to the problem of predicting the odds of events occurring within the range of an 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

Probability (event) $= \dfrac{1}{1 + e^{-Z}}$, 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} = total borrowings/total assets, X_2 = power and fuel/total income, X_9 = current liabilities and provisions/total assets and X_{13} = other fixed cost/total assets.The probability of the event not occurring is estimated as: probability (no event) = 1 − probability (event). The odd ratio i.e., the odd of being healthy is expressed as:

$$\text{odd} = \frac{\text{probability(event)}}{\text{probability(no event)}}$$

From the given data set with respect to one hundred companies, we have run the binary logistic regression with SAS package. As one knows, the estimation exercise with respect to the logistic regression 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 β_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 βs. 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 (λ) is defined as $\lambda = L_0/L_{max}$ where, L_0 = initial value of the like lihood function, and L_{max} = maximum of the same function. The appropriate test follows directly from the fact that $-2 \log \lambda = -2 \log \left[\frac{L_0}{L_{max}}\right]$ 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 βs for testing the reliability of the estimated βs. Next is the issue of goodness of fit. The goodness of fit in the 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) are used for testing the goodness of fit. The package also provides the Hosmer and Lemeshow goodness of fit. The results of the logit run are given in Table 8.

Reporting the results of binary logistic regression analysis

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

$$\mathrm{P}\,(\mathrm{Healthy}) = \frac{1}{1+e^{-Z}} \quad \text{where,}$$

$$Z = 9.5909 - 14.3800\,X_{19} - 46.9134\,X_2 - 15.5178\,X_9 + 126\,X_{13}$$

On the test of goodness of fit, Cox and Snell R-Square value (0.6988) and Max-rescaled R-Square value (0.9317) signify that our model is a good fit. 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 'percent concordant' (99.3%), 'Somers' D' (0.987), 'Gamma' (0.989) and a low value of 'percent discordant' (0.6) along with a low value of 'percent tied' (0.2) (Table 8). Result of testing global null hypothesis: BETA (β) = 0 shows that the chi-square value of the likelihood ratio statistic is significant even at 1% level of significance. This implies that 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% probability of being true. An analysis of the classification table generated from logistic regression output reveals that with probability of 0.72, 97% of the selected companies have been correctly classified into their respective predetermined group by our model. Out of 50 predetermined

Table 8 Logistic regression output.

Model information		
Response variable		Dependent
Number of response levels		2
Model		Binary logit
Optimisation technique		Fisher's scoring
Number of observations read		100
Number of observations used		100
Response Profile		
Ordered value	Dependent	Total frequency
1	1	50
2	0	50
Probability modelled is dependent = 1.		
Model convergence status:		
Convergence criterion (GCONV = 1E − 8) satisfied.		
Model fit statistics		
Criterion	Intercept only	Intercept and covariates
AIC	140.629	28.631
SC	143.235	41.657

Table 8 (*continued*)

Model fit statistics

Criterion	Intercept only	Intercept and covariates
−2 Log L	138.629	18.631

R-square: 0.6988.
Max-rescaled R-square: 0.9317.

Testing global null hypothesis: BETA = 0

Test	Chi-square	DF	Pr > Chi-sq
Likelihood ratio	119.9982	4	<0.0001
Score	71.9276	4	<0.0001
Wald	11.5553	4	0.0210

Analysis of maximum likelihood estimates

Parameter	DF	Estimate	Standard error	Wald chi-square	Pr > ChiSq
Intercept	1	9.5909	3.2883	8.5069	0.0035
X_2	1	−46.9134	17.1907	7.4475	0.0064
X_9	1	−15.5178	6.7776	5.2421	0.0220
X_{13}	1	126	55.7952	5.1014	0.0239
X_{19}	1	−14.3800	4.8673	8.7284	0.0031

Odds ratio estimates

Effect	Point estimate	95% Wald confidence limits	
X_2	<0.001	<0.001	<0.001
X_9	<0.001	<0.001	0.107
X_{13}	>999.999	>999.999	>999.999
X_{19}	<0.001	<0.001	0.008

Association of predicted probabilities and observed responses

Percent concordant	99.3	Somers' D	0.987
Percent discordant	0.6	Gamma	0.989
Percent tied	0.2	Tau-a	0.499
Pairs	2500	c	0.994

Partition for the Hosmer and Lemeshow test

Group	Total	Dependent = 1		Dependent = 0	
		Observed	Expected	Observed	Expected
1	10	0	0.00	10	10.00
2	10	0	0.00	10	10.00
3	10	0	0.01	10	9.99
4	10	1	0.11	9	9.89
5	10	0	1.72	10	8.28
6	10	9	8.30	1	1.70
7	10	10	9.87	0	0.13
8	10	10	9.98	0	0.02
9	9	9	9.00	0	0.00
10	11	11	11.00	0	0.00

Hosmer and Lemeshow goodness-of-fit test

Chi-square	DF	Pr > Chi-sq
9.7520	8	0.2829

healthy companies, the probability of only one company remaining as healthy is less than 1%. On the other side, out of fifty predetermined sick companies, probability of only six companies losing the status of sick varies from 29% to 54%. With respect to the estimated values of β coefficients, Wald Chi-square values of the estimators are satisfactory. Thus, with respect to β_0, 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 βs, i.e., β_1, β_2, β_3 and β_4, the estimated values are likely to be true with more than 95% probability. Expected signs of the estimated βs in our model are also realised. For example, sign of β_1, β_2 and β_3 is negative. The implication is that higher the value of X_{19}, X_2, and X_9, lower

Table 9 The classification matrix for holdout sample for the last four years.

Year	Predetermined group	No. of companies	Classified according to the model		Classification accuracy (percentage)
			Group		
			Healthy	Sick	
Four years before terminal year (TY)	Healthy	50	45	5	86
	Sick	50	9	41	
Three years before TY	Healthy	50	46	4	92
	Sick	50	4	46	
Two years before TY	Healthy	50	43	7	93
	Sick	50	0	50	
One year before TY	Healthy	50	45	5	95
	Sick	50	0	50	
Terminal year	Healthy	50	48	2	97
	Sick	50	1	49	

the odds that the event "healthy" will occur, i.e., lower is the probability that a company will become healthy. Finally, sign of β_4 is positive which implies that higher the value of X_{13}, higher are 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 expectations of a financial analyst.[12]

Validation of the model

We check the validity of our model by considering panel data of another one hundred companies from the PROWESS database comprising 50 healthy companies and 50 sick companies. The classification matrix (Table 9) shows that classification accuracy varies from 86% to 97%.

Summary and conclusion

Both the discriminant model and the predictive model we developed are based on accounting ratios having strong association with the macro indicators capable of classifying Indian manufacturing industries into "good performing" and "bad performing" at the macro level. Classification accuracy of the discriminant model is satisfactory; level of correct classification varies from 97% (one year before the terminal year) to 86% (four year prior to terminal year). The predictive model exhibits equally high levels of accuracy; with probability of 0.72, 97% of the selected companies are correctly classified into their respective predetermined groups. The models we have developed s are thus user-friendly, accessible to laymen, and would be useful in many respects. They would help investors in planning the buying or selling of stocks. Banks and financial institutions may use the models for evaluation of loan proposals, credit decisions and risk analyses. The predictive model would have wide application in turnaround management. Once a company shows fair probability of becoming sick in the near future, the management would be in a position to intervene immediately and take remedial measures for its turnaround. Similarly, lending institutions would be able to take a decision on debt restructuring of a company whose probability of getting sick is high.

Notwithstanding the robustness of our models and their application in various practical decision making situations, our study is not devoid of limitations. First, CMIE does not fully reveal how various items of balance sheet and profit and loss account of a company are clubbed to constitute the board categories of financial parameters, namely, net worth, shareholders' fund, term liabilities, short term bank borrowings, current liabilities, current assets, gross sales, net sales, cost of goods sold etc. In our study, we had to take these items as they are given in the PROWESS database. Despite our in-depth study of the various items of balance sheet and profit and loss account, we cannot hold that in the various items of the balance sheet and profit and loss account of the companies selected for our panel data would necessarily depict a true and fair picture. We should also point out that our study is based on financial data of Indian manufacturing companies in the organised sector and does not include companies in the small scale and unorganised sector. Units run by partnership firms, proprietorship firms, and registered cooperative societies have also not been studied by us. Since manufacturing units too fall within these excluded categories the validity of our models have thus not been tested on them in our study. The future research on the subject may take into account the above limitations.

[12] Suppose, for the year $Y_{t+1,}$ X_{19} is higher than X_{19} for the year Y_t. This may happen under three situations. First, value of total assets (TA) would remain same, but value of total borrowings (TB) would be more in $Y_{t+1.}$ Second, rate of increase in TB would be more than rate of increase in TA. Lastly, amount of TB would remain same, but TA for Y_{t+1} would be less than TA for Y_t. In any of these situations, possibility of a company's performance in Y_{t+1} would be bad compared to $Y_{t.}$ If such a trend continues, probability of the company getting sick is higher. Similar would be the position for X_9 and $X_{2.}$ On the other hand, higher the X_{13}, the value of sales and capacity utilisation is expected to be higher resulting in larger profit margin and the probability of the company getting sick would be less.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 45—69.

Gupta, L. C. (1993). *Financial ratios for monitoring corporate sickness*. Oxford University Press.

Pindyck, R. S., & Rubinfeld, D. L. (1985). *Econometric models and economic Forecasts*. McGraw Hill Book Company.