DOCTOR OF PHILOSOPHY

Multivariate analysis and survival analysis with application to company failure.

Shani, Najah Turki

Award date:
1991

Awarding institution:
Bangor University

MULTIVARIATE ANALYSIS AND SURVIVAL ANALYSIS WITH APPLICATION TO COMPANY FAILURE

A Thesis submitted to the University of Wales

BY

NAJAH TURKI SHANI , B.Sc. , M.Sc. , FSS

In candidature for the degree of

PHILOSOPHIAE DOCTOR

Centre for Applied Statistics,
School of Mathematics,
University College of North Wales,
Bangor, North Wales,
United Kingdom

1991

ACKNOWLEDGEMENTS

I take this opportunity to express my respectful appreciation of the contributions made by several persons towards the completion of this thesis.

First and foremost my supervisors, Dr. J. Y. Kassab and Professor S. Mcleay for their sustained interest, suggesting the topic, academic guidance and help with various difficulties which I have encountered during this study, all of which have contributed immensely to the quality of this research.

I appreciate the support given by the Centre for Applied statistics staff and the computer laboratory staff who offered valuable and helpful comments.

I would also like to thank Dr. J. Cummingham, the Deputy Chairman of the School of Mathematics, who was extremely helpful in providing facilities that were necessary for conducting this research and to the administrative staff of the Department.

I wish to thank those people, specially Mr. J. Curtis, who has given me the encouragement and moral support to persevere with my studies through to completion.

This research has been sponsored by the Higher Education Ministry of Iraq, whose financial assistance is gratefully acknowledged.

Finally, my warmest thanks go to my brother, Mr. S. T. Shani and my sister, Mrs. H. T. Shani for their financial support and encouragement.

SUMMARY

This thesis offers an explanation of the statistical modelling of corporate financial indicators in the context where the life of a company is terminated. Whilst it is natural for companies to fail or close down, an excess of failure causes a reduction in the activity of the economy as a whole. Therefore, studies on business failure identification leading to models which may provide early warnings of impending financial crisis may make some contribution to improving economic welfare. This study considers a number of bankruptcy prediction models such as multiple discriminant analysis and logit, and then introduces survival analysis as a means of modelling corporate failure. Then, with a data set of UK companies which failed, or were taken over, or were still operating when the information was collected, we provide estimates of failure probabilities as a function of survival time, and we specify the significance of financial characteristics which are covariates of survival. Three innovative statistical methods are introduced. First, a likelihood solution is provided to the problem of takeovers and mergers in order to incorporate such events into the dichotomous outcome of failure and survival. Second, we move away from the more conventional matched pairs sampling framework to one that reflects the prior probabilities of failure and construct a sample of observations which are randomly censored, using stratified sampling to reflect the structure of the group of failed companies. The third innovation concerns the specification of survival models, which relate the hazard function to the length of survival time and to a set of financial ratios as predictors. These models also provide estimates of the rate of failure and of the parameters of the survival function. The overall adequacy of these models has been assessed using residual analysis and it has been found that the Weibull regression model fitted the data better than other parametric models. The proportional hazard model also fitted the data adequately and appears to provide a promising approach to the prediction of financial distress. Finally, the empirical analysis reported in this thesis suggests that survival models have lower classification error than discriminant and logit models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER ONE

INTRODUCTION

This thesis is concerned with the statistical modelling of corporate financial indicators in the context where the life of a company is terminated. Whilst it is natural for companies to fail or close down, as they do for various reasons (see Section 2.3), an excess of failures cause a reduction in the economic activity of the economy as a whole. These failures not only reduce government income by both lost revenue from taxation and a negative flow of funds through social security payments, they also affect the economic well-being of other businesses which lose sales and of investors who lose income and capital. There are occasions when, for either political or economic reasons, the government will decide to prevent the failure of a large company important to the state of the nation (for example the British government support of the failing motor industries in the 1970's). Perhaps, if these problems had been identified at an earlier stage then government financial support may not have been so costly. Therefore, studies on business failure identification leading to models which may provide early warnings of impending financial crisis may make some contribution to improving economic welfare. Rather than waiting until the event of failure to diagnose the problem, the emphasis should be on identifying failing companies during the early stages of their decline. Hopefully, some corrective action could be taken to stem the decline. For the past 20 years, various business administration specialists have presented articles in financial journals that have employed multivariate statistical techniques on corporate financial data in order to develop statistical models which would identify failing companies. These models predicted failure with a high degree of accuracy when companies were near collapse. As failure becomes more remote in time, however, the forecasting accuracy of these models declined. Some of these models also contained statistical weaknesses that cast doubt on their results. This thesis considers a number of bankruptcy prediction models, and then introduces survival analysis as a means of modelling corporate failure. In this context, we look upon company development over time as a process in which companies are created and, eventually, are taken over or fail. At any given point in time, the survivors are those companies which have not yet been taken over or have not yet failed. Essentially, the approach taken in this thesis is to introduce company financial characteristics into a probability model as covariates of survival, treating surviving companies as censored observations whose eventual fate is unknown.

One can view the development in certain financial ratios as indicators of the company's state of health, although failure may be caused by different circumstances. Beaver (1966), Altman (1968) and, subsequently, many other authors have carried out research in this field. Their evidence indicates clearly that with a few financial measures (e.g. financial ratios) corporate failure can be predicted for a period of at least five years before failure. Naturally, it is possible that failure prediction models might benefit from the inclusion of other financial and non-financial variables. However, the data used here is restricted to a broad set of financial ratios as used conventionally in many studies and the focus of this thesis is on the statistical methodology and its applicability to the phenomenon under investigation, i.e. company failure in a broad sense.

1.1 <u>Objectives of the Study</u>

The preceding discussion justifies the development of a warning system for financial distress in a business company. The objectives of this thesis are to construct statistical models that can identify in advance those companies that will become financially distressed and an attempt to understand the structure of large cross-sectional accounting information sets. For this thesis a data set relating to financial statement information of large U.K. industrial companies was gathered from EXSTAT. This is an extensive data base that has also been analyzed by others in building failure prediction models, and in understanding the structure of large cross-sectional accounting information sets. Then, with the specific data set compiled for this study of U.K. companies which failed, or which were taken over, or which were still operating when the information was collected, we provide estimates of failure probabilities as a function of survival time, and we specify the significance of financial characteristics which are covariates of survival. There is some evidence of survival bias in the time series of certain financial ratios, after treating financial disclosures as non-synchronous, irregular repeated measures when estimating mean effects. However, survival bias is not the central issue. The core of this thesis is concerned with three innovative statistical methods. First, a likelihood solution is provided to the problem of takeovers and mergers in order to incorporate such events into the dichotomous outcome of failure and survival. Second is the specification of a parametric and non-parametric model of company survival, where we evaluate the assumptions of the model on the basis of an analysis of residuals, and select between Weibull, Exponential and Log-logistic regression models for best fit and accuracy of prediction. Third, a "randomly-censored stratified sampling" solution is provided to the problem raised by moving from a matched sampling basis to one where the structure of the survivor group no longer reflects that of non-survivors.

1.2 Chapter survey

This thesis is divided into seven Chapters each one presenting a different phase of the study :

Chapter one introduction

Chapter two is a brief survey of the literature of business failure prediction. Included are definitions of business failure and contrasting authors' views on the causes of failure. This chapter concludes with a survey of the major models that have been used in published studies. Of particular importance are the statistical techniques used, and the overall forecast accuracy of each, and some consideration is given to the variables in the models.

Chapter three contains an extensive discussion of the construction of the new data set used for this study. Accounting data on 463 companies was collected from the EXSTAT source, restricted to industrial companies. The 463 companies consisted of companies which went into liquidation ("bankrupt"), companies which combined with others or were acquired by others ("merged") some which closed down for other reasons or moved from the U.K. ("other") and surviving industrial companies. The methods used to identify non-surviving companies and to collect data are explained. Also presented in this chapter is some exploratory analysis of the general time series behavior of financial ratios of these companies, a discussion of the computer data analysis used to study the distributional properties of the data, and is concerned with the application of principal component analysis, which is used for structural simplification so that the large number of variables may be reduced to fewer components.

Chapter four is concerned with the application of stepwise discriminant analysis and quadratic discriminant analysis, which are used to determine the most important financial ratios that are associated with the failure of a company and to predict the probabilities of failure, first, *before* reclassifying "merged" and "other" companies and, second, *after* reclassifying "merged" and "other" companies. Also presented are methods used to reclassify "merged" and "other" companies into either the "bankrupt" category or the "live" category, including stepwise discriminant analysis and survival analysis based on the Weibull model. Each of the two methods are carefully explained and the results of the two methods are contrasted.

Chapter five is concerned with the application of the logistic model, which is used to predict failure and to determine for the data under investigation the most important financial ratios affecting the outcome. The explanatory variables used in the prediction models are identified by stepwise regression.

Chapter six considers the covariates of survival which are modelled in an attempt to understand the structure of the large cross-sectional accounting information set under investigation. The models used are based on the hazard function. Two classes of such models are considered: parametric models which contain Weibull, Exponential and Log-logistic regression models, and the non-parametric proportional hazard model. Parameter estimation is based on maximum likelihood estimation.

Chapter seven contains the general conclusions of the study, and identifies the potential contributions of the statistical modelling approach to applications in financial analysis.

CHAPTER TWO

BUSINESS FAILURE PREDICTION MODELS

2.1 Business Failure

Necessary to any statistical model of corporate failure are certain basic inputs. First, a definition is needed as to what exactly is a "business failure". Second, financial writers (e.g. Lev, 1974, and Dewing, 1941) have suggested the causes and warning signs of failure- how could a model be built which would associate failure with these signs? Other items that must be considered are: existing failure prediction models, the success and limitations of these models, and new theoretical techniques that offer a solution to the weaknesses of existing models.

2.2 Business Failure Definition

There are many institutional aspects of corporate failure that figure indirectly in model building. These have been studied by many experts in various fields. Economists study the effects of national policy decisions on business and the costs of failure on the economy. Financial experts are concerned with the loss of investment in failed firms. Legal experts argue over the payment of creditors' claims. These professionals are all concerned with determining the costs of business failure and who pays these costs.

Various definitions of business failure have been presented by different authors.

Beaver (1966), defines failure as the inability of a firm to pay its financial obligations as they mature. Operationally, a firm is said to have failed when any of the following events have occurred : bankruptcy , bond default , an overdrawn bank account, or nonpayment of a preferred stock dividend.

Altman (1971), defines economic failure by economic criteria, where the realized rate of return on invested capital, after allowing for risk, is significantly and continually lower than prevailing rates on similar investments.

Deakin (1972), defines failure to include only those firms which experienced bankruptcy, insolvency, or were otherwise liquidated for the benefit of creditors.

Taffler (1982), defined the failure as receivership, voluntary liquidation (creditors),or winding up by court order.

Thus, in the context of failure prediction, the concept of failure varies from (i) the broad definition of a company which is unable to settle its financial obligations (which may be a temporary state of affairs resolved by a reorganization of financial structure) to (ii) the narrower definition where the company is liquidated.

The first can be considered as "technical insolvency" which refers to the inability of a firm to meet its currently maturing obligations (Walter, 1957). It may be only a temporary condition for the firm. For instance, the firm may have a positive equity position and a sufficiently good outlook to get short run financial help over its present cash crisis. On, the other hand, when a firm is in such a bad position that it cannot pay its debts and secure new financing, then it can voluntarily or involuntarily enter into bankruptcy. This leads to an alternative conception of the "life" and "death" of companies. At any one point in time t, we may observe companies which are in existence. To that stock of companies will be added newly-created companies, and there will also be companies which close down in the intervening period and, therefore, do not survive until t+1. However, it is not necessarily the case that a firm which closes down does so because it has failed. For instance, one company may be acquired by another, although there exists the possibility that the takeover target was heading for failure and its restructuring is effected by absorption into another company. Therefore, in the same way as other researchers have attempted in the past to broaden the definition of failure to encompass temporarily bankrupt firms (i.e. technically insolvent), so in this thesis the partition is between "survivors" and "non-survivors", with the latter group comprising (i) failures and (ii) companies which closed down for other reasons and which we may wish to partition between failed non-survivors and other non-survivors.

2.3 Causes of Failure

The ability to predict corporate failure is important for all parties involved in the corporation, in particular for management and investors. An early warning signal of probable failure may enable them to take preventative measures: changes in operating policy or reorganization of financial structure, but also voluntary liquidation could shorten the period over which losses are incurred. The possibility of predicting failure is important also from a social point of view, because such an event may be an indication of misallocation of resources; prediction provides opportunities to take corrective measures.

No Theory Of Corporate Failure

Since the objective of this study is to develop a quantitative model to predict corporate failure, a generally accepted theory of corporate failure is the place to start in formulating a model. A survey of the finance literature reveals that there is no well-formulated theory of corporate failure, and Lev (1974), gives the following reasons for this:

1. the complexity and diversity of business operations
2. the lack of a well-defined economic theory of the firm under uncertainty, and
3. the reluctance of theorists to study failure and include it in their models.

Because of this lack of theory, model builders have considered the reasons that financial experts have suggested as being the explanations for the financial decline of firms.

The idea of bad management is perhaps the most mentioned cause of business failure. Dewing (1941), the author of a classic text in corporate finance, wrote: "The usual causes assigned for failure are, in truth, not causes but excuses; the real cause is the lack of those human qualities which, for want of a better understanding of the human mind, we epitomize by the expression management. Unfortunately, bad management is not a readily identifiable and quantifiable item. Further, corporations are not required to disclose decisions made by their top managers. As an observer of corporations, we can only note the later effects of management decisions".

Dewing (1941) also lists four fundamental economic causes of failure:

1. excessive competition
2. unprofitable expansion
3. change in public demand for the commodity
4. the distribution of capital as ostensible profit.

Dewing considers the second reason, unprofitable expansion, as the prime reason for failure. Dewing makes this point as a result of observing business failures in the 1920's and 1930's.

Financial writers have always considered the state of the economy as an important element in the financial health of firms. For instance, Altman (1971) developed a regression model showing that change in corporate failure rate is inversely associated with changes in GNP, stock prices, and money supply. Gordon (1971), in an article on financial distress of corporations, notes that when corporate debt and interest payments are at record levels, and the government committed to an anti-inflation policy, the likelihood of failure is increased. For a firm with operating losses plus high leverage, cash runs out, new credit is not available and The firm fails. In periods of economic slowdown, the number of business failures increases.

2.4 Survey of Failure Prediction Models

2.4.1 Corporate Failure Prediction Model

Failure prediction models can be of help to investors in debt securities when assessing the likelihood of a company experiencing problems in paying interest or principal repayments. Also, the failure of a business firm is an event which can produce significant losses to creditors and stockholders. Therefore a model which predicts potential business failure as early as possible would help to reduce such losses by providing sufficient warning. The predictive value of financial ratios and related financial data has received considerable attention in recent years.

This was a sufficient motivation for Beaver (1966) and Altman (1968) to develop models for predicting failure based on the financial information disclosed by firms. Research conducted more recently on the use of financial ratios to predict failure can be divided into two groups, the univariate and the multivariate studies. The first group is concerned with the predictive ability of individual financial ratios; whereas the second group in more concerned with performance of several ratios combined together to predict failure.

2.4.1.1 Univariate Models

A univariate model to predict financial failure involves the use of a single variable in a prediction model. There are two assumptions in this approach (Foster, 1986):

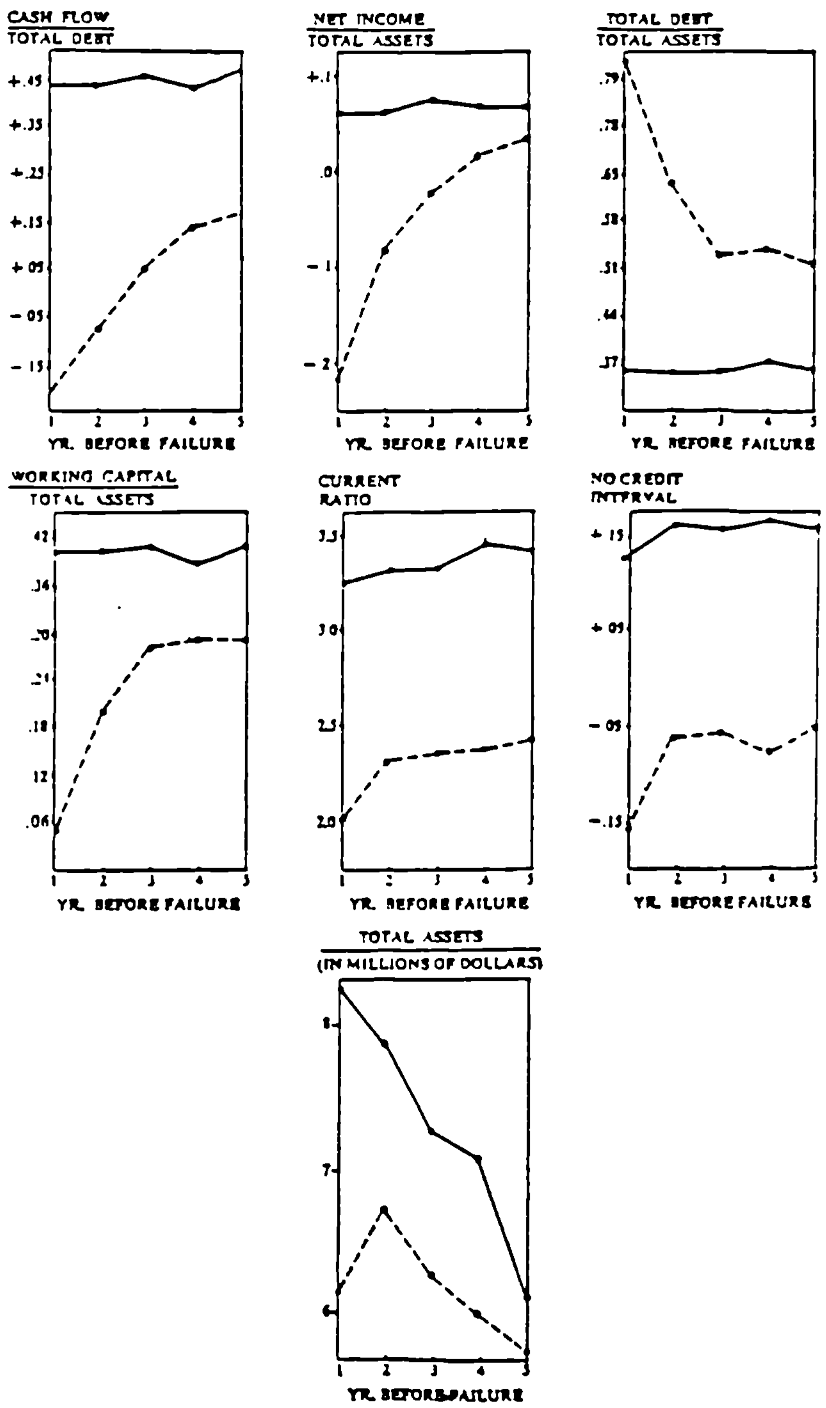
1. the distribution of a variable for the distressed firms differs systematically from the distribution of the variable for the non-distressed firms
2. this systematic distribution difference can be utilized for prediction purposes.

Beaver's model

In his seminal paper, Beaver (1966) developed and tested a univariate prediction model of corporate failure based on observations of 79 failed firms which were each matched for industry classification and asset size with another firm that continued in operation over the period 1954-1964. The effect of this pairing was to stratify for size and industry factors. Although the study's results were therefore only applicable to that stratum of firms, failures amongst the groups in question were more serious than elsewhere, and data was more available for these firms. Financial statement data for five years prior to failure was examined. Some thirty different ratios were selected among ratio groups, that were believed to be important (profitability, leverage, activity, cash flow, etc.) . Mean values for each variable over five years for failed and non-failed firms were examined. The mean ratios of the failed firms show distinct trend. Whereas the mean ratios of the non-failed firms remained relatively constant. Figure

2.1 shows that The trend in the mean ratios was very pronounced for the failed firms over the five year period prior to failure.

Figure 2.1 Profile analysis, comparison of mean values from Beaver (1966).



Legend: —— Non-failed firms ---- Failed firms.

Differences were said to offer proof of the ability of financial ratios to predict failure. Then one ratio was selected, cash flow to total debt, as the single best predictor of failure, because it had the smallest classification error (i.e., 13%) in discriminating failed from non-failed firms.

Although these findings suggested that ratio analysis could be useful for as early as five years before failure, Beaver cautioned that ratios have to be used selectively. He found that not all ratios predict equally well. Further more, different ratios do not predict failed and non-failed firms with the same degree of success.

Since the univariate methodology places emphasis on only individual signals, it is possible that different financial ratios might provide conflicting signals of firm's financial condition. Thus, Altman (1968) commented that Beaver's approach to ratio analysis " is susceptible to faulty interpretation and is potentially confusing ", and suggested using multivariate analysis to investigate the predictive ability of financial ratios.

2.4.1.2 Multivariate Models

A multivariate model for predicting financial failure involves the use of several variables in a prediction equation. Multivariate models of financial distress have been developed in various countries including the United States, Japan, Germany, Switzerland, Brazil, Australia, England, Canada, the Netherlands and France as described in Altman (1984), An International Survey.

Altman's model

The first multiple discriminant analysis (MDA) model was published in September 1968 by Edward Altman. He developed a model for the prediction of corporate bankruptcy on a sample of 33 bankrupt and 33 non-bankrupt firms over the period 1946-1965. The 33 bankrupt manufacturing firms had filed a bankruptcy partition under Chapter X of the Bankruptcy Act. For each of the bankrupt firms a comparable match was chosen from the same industry with similar assets size measured over the same chronological period. The asset size of bankrupt firms was from 0.7 to 25.9 million dollars one year prior to bankruptcy. 22 ratios were selected based on their popularity in the literature and their potential relevance to the study. These ratios were classified into 5 standard categories -- liquidity, profitability, leverage, solvency and activity. He used many computer runs to select the best five variables out of 22 financial ratios. His model is:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

where:

X_1 = Working capital to total assets

X_2 = Retained earnings to total assets

X_3 = Earnings before interest and taxes to total assets

X_4 = Market value of equity to book value of total debt

X_5 = Sales to total assets

Z = Represents the discriminant score of the firm.

Altman classifies firms with Z scores as follows:

greater than 2.99 - non-bankrupt

less than 1.81 - bankrupt

between 1.81 and 2.99 - "zone of ignorance" or "gray area".

At the beginning of his paper Altman states that he wants to examine the usefulness of financial ratios in a model to corporate bankruptcy. After presenting his model, he tested his model on various data sets. His conclusions are that his model correctly classifies firms in years before bankruptcy as follows:

95% one year before

72% two years before

48% three years before

29% four years before

36% five years before.

Thus, the model proved to be an accurate forecaster of failure when failure is imminent. As failure becomes more remote, predictive accuracy drops. This decline in predictive accuracy is another important aspect of the failure prediction problem.

Thus, by developing a linear discriminant function which combined five financial ratios, Altman obtained an approach that out performed Beaver's "cash flow to total debt" method in predicting firm's failure. Altman's study is considered as the pioneering research in applying multivariate techniques to develop a predictive procedure using financial ratios.

Limitation of Altman's (1968) model include the following:

1. One limitation is that of the ex-post nature of the analysis, i.e. the estimation and validation samples both include firms that are known to have failed on a set date. Thus, it is possible in the research to compare the financial ratios of failed and non-failed firms one year, two years, etc., prior to failure. However, in decision-making contexts, one knows neither which firms will fail nor the date on which they fail. To demonstrate that the results of

this research have direct applicability to decision contexts, it would be necessary to make ex-ante predictions about the failure (and its timing) of firms currently non failed.

2. Little effort was directed towards the construction of a testable theory that would specify the variables to be included in the discriminant function. The approach of choosing 22 variables and then using a stepwise discriminant model to select the variables in the final discriminant function may be limited in its ability to provide generalizable results as to what financial variables are likely to be consistent predictors of financial distress.
3. The sample of firms used by Altman consisted of matched samples of bankrupt and non-bankrupt firms, selected on a non-random basis. However, no additional procedures were used to overcome the limitations of having a non-random selection of the original sample. The result is that the parameters estimated would be subject to bias; some characteristics may be over-represented in the samples. Thus, the resulting discriminant function may be sample specific.
4. The use of a paired-sample design where firms are matched on size and industry criteria effectively precludes these variables as indicators of financial distress in the study. There is considerable evidence that both size and industry groups contain information on distress likelihood.
5. The use of equal-sized samples of bankrupt and non-bankrupt firms also distorts the actual prior probabilities of firm's belonging to either group. Deakins (1977) analysis of this type of bias indicates that under such circumstances, the stated error rates may not reflect the extent of each type of error. The most serious

effect would be a tendency to understate the misclassification of non-failing companies into the failing group.

6. Altman's method does not ~~depends~~ depend on the additional assumption that the variables distributed as the multivariate normal distribution. 7

The use of linear discriminant analysis assumes that the data for the failed and non-failed firms have the same dispersion matrix. In a later study, Altman, et al (1977) tested this assumption and found that the dispersion matrices of the failed and non-failed firms could not be considered identical. No test of multivariate normality was conducted. It was determined that a quadratic discriminant analysis was required. Stepwise exclusion was used to limit the twenty-seven variable set to seven discriminatory variables. The results indicate that even though the quadratic classifier is statistically more appropriate, the linear classifier gives a lower classification error rate. The holdout sample used was not an independent sample, since data from the original sample's financial statements 2-5 years prior to failure are applied to the parameters established from one-year prior data, suggesting that a comparison of these results with other studies is not appropriate.

Altman (1984), reviewed and compared a relatively large number of empirical failure classification models from 10 countries. Much of the material is derived from little known or unknown sources. Indeed as financial institutions and government agencies in various countries, e.g., Canada, U.S., Brazil, France, and England, wrestle with the problem of large firm failures in the future, the knowledge that prior work has been done with respect to early warning models may help avoid the consequences or reduce the number of these failures. In concluding, Altman notes that he expects the quality and reliability of models constructed in many of the reviewed countries to improve as:

1. the quality of information on companies is expanded and refined
2. the number of business failures increase thereby providing more data points for empirical analysis, and
3. as researchers and practitioners become more aware of the problems and potential of such models.

Where sufficient data does not exist for specific sector models, e.g, manufacturing, retailing, and service firms, the application of industry relative measures can perhaps provide a satisfactory framework for meaningful analysis.

Deakin's model

Deakin (1972) applied multivaiate discriminant analysis (MDA) to 14 financial ratios initially used by Beaver (1966). His estimation set consisted of 32 pairs of firm - bankrupt matched with non-bankrupt over the period 1964-1970. He obtained a liner discriminant function in which all the 14 variables were found to contribute significantly to the discriminating ability of the function. In general, his discriminant function was able to predict business failure as far as three years in a advance with an accuracy of around 94%. Rather than using a critical value for classifying the cases, Deakin used a modification of discriminant analysis that assigns probabilities for membership to the classes. Each firm was reclassified each year in a manner that weighted the probability of group membership with its deviation scores from prior periods. This technique improved the classification error rate significantly over those found by either Beaver or Altman. Using such probability estimates for group membership, Deakin's model provided error levels of 3 percent, 4.5 percent, and 4.5 percent respectively for the first three years prior to failure. The error rates for the fourth and fifth years increased to 21 percent and 17 percent respectively. These results appear to be an improvement over Altman's, which could only predict accurately in the first year prior to bankruptcy. Comparison with other studies is difficult because the choice of a critical probability for assignment to a group is a subjective choice even though the technique will generate probability of group membership. Ideally such a choice should be based on an analysis of cost of errors. No such analysis was reported in Deakin's study. Also the method of group membership assignment according to probabilities was not discussed. The limitation of Altman's model cited apply to Deakin's model as well. Deakin's method depends on the additional assumption that the variables were distributed as the multivariate normal distribution, but no multivariate normality test was provided.

A later study by Deakin (1976) also found that financial ratios were non-normal. Since univariate normality is a necessary but not sufficient condition for the normality of these variable's joint distribution, the adherence to the assumption of multivariate normality is doubtful. Lack of adherence to these assumption could affect the predictive results.

<u>Blum's model</u>

Blum (1974) developed a model for predicting failure using a sample of 115 non-failed firms over the period 1954-1968. For each of the failed firms, a similar non-failed firm was chosen from the same industry, size and fiscal year. In Blum's model the accuracy of the failing company model in distinguishing failing from non-failing firms was tested by using discriminant analysis for computing an index and a cutoff point on the index. The index is derived from the financial model by computing the values of each of its variables for each company studied. When the variables for one company are standardized and added together, their sum is that company's index score. A critical score exists which results in a minimum of misclassification. If all companies with index scores above the critical score are predicted to succeed and all companies with scores below are predicted to fail, erroneous predictions will be minimized. When a firm with an unknown group identity (failed-nonfailed) is classified by a discriminant function as similar to firms which failed in the next year, the firm's classification will be treated as a prediction that the firm will fail one year from the date of prediction. However, his model contained three new features:

1. The financial ratios included in the discriminant function of the earlier studies were selected on the basis of either their popularity, subjective judgments by the research, or the result of an elimination process using stepwise regression (as in Edmister's, 1972). In contrast Blum constructed a "cash flow framework" to theoretically identify the factors that will affect the probability of failure. Ratios associated with these factors were used as explanatory variables in the discriminant function.
2. Blum identified three groups of relevant factors through his "cash flow framework". The first group consisted of liquidity-related ratios.The second group had only one factor: the rate of return to common stockholders. The third group consisted of measures of variability of income and the net quick assets to inventory ratio over a time period. None of the third group of factors had been

used as an explanatory variable in previous studies.

3. Earlier studies investigated the change in accuracy when multivariate discriminant analysis is used to predict failure for different lengths of time ahead of failure. They all reached the intuitively obvious conclusion that predictions become more accurate the closer one gets to the actual date of failure. Besides studying this effect, Blum also studied the effect of using different numbers of years of prior data to predict failure for a given time period ahead. His primary purpose was to investigate the number of years of prior data required to improve predictive accuracy. His findings suggested that for predicting one year ahead using more than six years of data would actually reduce instead of increase the predictive accuracy of the resultant discriminant function.

Blum's empirical results indicated that the discriminant functions from his "failing company model" could achieve 93-95% accuracy for predictions one year before actual failure, 80% accuracy at second year before failure, and 70% accuracy at third, fourth and fifth years before failure. Blum's model was used for legal decisions in the U.S.A., the so called 'failing company doctrine' is used as one defence against an antitrust law in the U.S.A.. This doctrine can apply where one of two merging companies is likely to fail and where the failing company has received no offer to merge from a company with which a merger would have been legal. Predicting failure using Blum's model provides the court with some evidence as to which firms may lay within the failing company doctrine defence against antitrust laws. His model's performance compared favourably with the results obtained by Deakin (1972), Altman (1968) and Beaver (1966).

Mason and Harris model

Mason and Harris (1979) developed a model specifically for the identification of construction companies in danger of failure. The study was carried out because of the concern that in the U.K., at least, contracts often tend to be awarded on the basis of price without adequate consideration of contractor's solvency and thus his ability to complete the work. 20 construction companies failure between 1969 and 1978 constituted the failed set and the continuing sample consisted of 20 particularly sound concerns on a traditional financial ratio analysis basis with 1976-1977 accounts used. A list of 28 discriminant variables was developed using a stepwise linear discriminant analysis by finding the variable that discriminates most between the groups of known "failed" and "solvent" companies. It then combines this variable with each of the other variables in turn until it finds the variable which contributes most to any further discrimination of groups and then continues in a similar manner until very little discrimination is gained by inclusion of a further variable. The following model was derived:

$$Z = 25.4 - 51.2X_1 + 87.8X_2 - 4.8X_3 - 14.5X_4 - 9.1X_5 - 4.5X_6$$

where:

X_1 = profit before interest and tax to opening net assets

X_2 = profit before interest and tax to opening net capital

X_3 = debtors to creditors

X_4 = current liabilities to current assets

X_5 = $\log_{10}$(days debtors)

X_6 = creditors trend measurement.

None of the 40 firms was misclassified but there were 4 type I errors in a validation sample of 11 failed enterprises (36.3%). Also 58% of the total 31 failed enterprises had failing characteristics 4 years before failure. Mason and Harris also used the Bayesian statistical approach to find an indication of how many "at risk" companies will fail each year considering the following Bayesian formula:

$$P(F/C_r) = \frac{P(C_r/F)}{P(C_r/F) + P(C_r/NF)}$$

where

$P(F/C_r)$ is the probability of a company, classified by the model as "failed", actually failing

$P(C_r/F)$ is the probability of a company being insolvent and also being classified as "failed"

$P(C_r/NF)$ is the probability of a company being classified as "failed", but being actually solvent.

Therefore, 18% of the companies classified by the model as "failed" should actually fail each year. However, their model is not only able to distinguish between known failed and solvent companies on a historic basis, but that it has "true" predictive ability in a statistical sense. It has been shown that the model is able to identify a short list of companies that are "at risk" of failure, and that it is also able to give an indication of the proportion of these firms that are likely to fail in the near future. Mason and Harris did not try to investigate the distribution of the variables before using the discriminant analysis.

Ohlson's model

Due to the restrictive assumption of discriminant analysis, Ohlson (1980) used the conditional logit analysis to construct his probabilistic bankruptcy prediction model. This statistical method avoids most of the disadvantages of discriminant analysis; the requirement concerning the distributional properties of the ratios, the output from the model which is a score not a probability of failure and the problems arising from the use of matched samples. Ohlson's model assumes that $P(X_i, \beta)$ is the probability of bankruptcy for any given X_i and β (where X_i is the predictor variable and β is unknown parameters). P is some probability function, $0 \le P \le 1$. The logistic function is

$$P = (1 + \exp(-Y_i)^{-1})$$

where $Y_i = \sum_j \beta_j X_{ij}$.

The use of logit means that no assumptions have to be made regarding prior probabilities of bankruptcy and/or the distribution of predictors. Ohlson also abandoned the use of a matched sample. Nine independent variables were selected. Firm size was included as a variable, calculated as log(total assets/GNP price-level index). Total asset size was also used to standardize three of the other variables, and current assets were used to standardize a fourth variable. Ohlson adjusted the firm size variable for price level changes in order to allow "real time implementation of the model", but it was the only variable adjusted in the set of nine. The sample of failed firms was selected from the Wall Street Journal Index. The firms included had failed between 1970 and 1976, they were industrials and had to have been traded on the stock exchange for at least three years prior to failure. Firms that did not report funds statements for the entire sample period were eliminated, leaving a sample of 105 bankrupt firms. In the non-failed sample, each of 2058 non-failed industrial firms was allowed to contribute one year of data to the data used in estimating the models. This means that no matching procedure was used, allowing the non-failed set to be a random sample. Three models were estimated, the first to predict failure within one year, the second to predict failure within two years if the firm did not fail in the first year, and the third to predict failure in one to two years. The coefficients of seven variables were found to be significant at at least the 0.10 significance level. The size variable was found significant at the 0.01 significance level in all three models. Other significant variables were total liabilities/total assets, working capital/total assets, net income/ total assets, funds from operations/total assets, and a dummy variable representing negative owner's equity. Classification errors were evaluated using the same set of data from which the model were estimated. Ohlson used this procedure for four reasons. First, he did not see his objective as "getting a precise evaluation of a predictive model". Second no "data dredging" was used to find a superior model. Third, unlike discriminant analysis, the logit technique is not an optimizing model. Fourth, the sample size is large, which would reduce the bias stemming from the lack of using a holdout sample. Assuming that the effects of Type I and Type II error rates are additive and that the best model minimizes the total error rate, a critical probability for classification was selected as 0.038. Thus if a firm's predicted probability of non-failure was below 0.038 the firm was classified as failing . Using this classification procedure for the first model, the misclassification rates were 17.4 percent for the non-bankrupt firms and 12.4 percent for the bankrupt firms in the first year prior to the failure date, which is significantly higher than those achieved by discriminant analysis studies. Error rates for the second and third year models were not reported. Since there are statistical reasons for believing that the logit technique can improve on the results from discriminant analysis, Ohlson found the results from his study disappointing.

The restrictive assumptions of discriminant analysis were not required, and interpretation of individual coefficients is appropriate in the logit model. This model lends itself, therefore, to broader research applications than discriminant analysis models.

The failure of Ohlson's model to achieve accurate predictions indicates that further refinements are necessary.

Taffler's model

Taffler's study (1982) used industrial enterprises quoted on the London Stock Exchange. The failed set of 23 firms consisted of all those companies failing between 1968 and 1973 and meeting certain criteria to ensure data completeness, consistency and reliability. Failure (bankruptcy) was defined as receivership, voluntary liquidation (creditors), winding up by court order or equivalent. The sample of non-failed firms was constructed differently to previous studies in that no matching with failed firms by industry, size or financial year was attempted nor was the number of firms made equal. Taffler argued that restricting the size of the non-failed sample to that of the failed sample only serves to restrict the total sample size and degrees of freedom, because the degree of freedom depends on the sample size. There is no point in restricting the sample size to match that of the failed companies, this simply reduces the size of the non-failed companies with consequent reduction in degrees of freedom. He also argued that in order to make valid inferences it is necessary for the sample groups employed in the analysis to be representative of their underlying populations, and he therefore suggested that the matching of continuing firms with failed firms by industry is incorrect since this does not provide for the non-failed set to be a random selection of all presently continuing industrial firms, particularly as some industries are more failure-prone than others. The same goes for attempted matching both by company size and financial year. Taffler's data for the non-failed firms were drawn from financial statements where financial year ends were in the calendar years 1972-1973.

The set of 45 non-failed firms was finally obtained meeting the initial industry, data availability and consistency requirements and most importantly that the firms must be financially sound. This is the most important departure from other studies in the selection of non-failed firms. Taffler explicitly recognized that a continuing firm is not necessarily financially healthy and that many companies presently in existence closely resemble previous bankrupts in terms of their financial characteristics. Three classes of discriminant variable were developed: conventional ratios, 4-year trend measures and fund statement variables. He found that the latter were too volatile for meaningful analysis and the trend measures added very little to the power of the discriminant model. Taffler therefore focused his analysis on a set of 50 financial ratios. The distribution of the straight ratio and trend measures were transformed (logarithmic or reciprocal) where appropriate to improve normality. They were then winsorized with any outliers beyond four standard deviations (s), from the mean of the remaining observations replaced by the mean and those between 2.5s and 4s by the appropriate 2.5s limit. Those variables which remained highly non-normal with many extreme values were omitted from further analysis. using a stepwise liner discriminant analysis, a model consisting of the following five variables was produced:

X_1 = Earnings before interest and tax to opening total assets

X_2 = Total liabilities to net capital employed

X_3 = Quick assets to total assets

X_4 = Working capital to net worth

X_5 = Stockturn.

Application of the z-model to the failed sample for prior years showed that nine of the 23 firms appeared sound on the basis of their penultimate accounts and only eight having failure characteristics 4 years before failure. Taffler has tested his model for its predictability by applying the model to 33 quoted manufacturing firms identified as going bankrupt between 1974 and 1976. He carried on to argue that a conservative estimate of the annual failure rate for the period 1974-1976 would be at least 2.5 percent. Taffler used the Bayes' theorem by letting F denote the event failure in the next year, AR a current at risk z-score and $\overline{AR}$ a current solvent z-score by considering the following Bayesian formula:

$$P(F/AR) = \frac{P(AR/F)\ P(F)}{P(AR)}$$

and

$$P(F/\overline{AR}) = \frac{P(\overline{AR}/F)\ P(F)}{P(\overline{AR})} .$$

Using Bayes' theorem taking his type I error of 12.1 percent and the 10.7 percent of companies with at risk scores, he suggested that the probability of failure given an at risk profile in the next year was 20.5% and the equivalent figure gives a financially healthy z-score was 0.34%. Taffler redeveloped his model and tested it on an ex-ante (i.e. forecast the future state of a firm given present data) basis. However, the results proved disappointing with 40.4% of the 52 failed firms being misclassified by the model.

Betts and Belhoul model

Betts and Belhoul (1982) develop a z-model consisting of the following variables, in terms of importance :

X_1 = profit before interest and tax to total assets

X_2 = quick assets to current assets

X_4 = current assets to net capital employed

X_3 = working capital to net worth

X_5 = days creditors

The two samples were 26 quoted companies failing mainly between 1974-1977 and 131 'going concerns' sampled randomly from the EXSTAT tape. A set of 26 potentially discriminating financial ratios were derived for the two groups and a conventional stepwise linear discriminant approach adopted to derive the model. No type I errors were registered and only 5 type II. Applying the model to an end 1979 EXSTAT tape led to only 6.1% of the 1230 enterprises registering a failing profile which the authors considered to be on the low side. There were 5 type I errors in a validation sample of 22 recent failures.

Zmijewski's model

Zmijewski (1984) examined the problems with non-random sample selection in models of financial distress. He points out that "Researchers typically estimate financial distress prediction models on non-random samples. Estimating models on such samples can result in biased parameter and probability estimates if appropriate estimation techniques are not used". He discusses two processes by which the random selection criterion may be violated, namely the choice based sample bias and the sample selection bias. The choice based sample selection bias arises because of the raw frequency rate of firms exhibiting financial distress characteristic.

Zmijewski points to many studies which have used the paired sample design (e.g. Altman, 1968) and concludes that "these studies estimated models on non-random samples which have compositions considerably from the population's composition". This violates the random selection assumption and he feels that the "dependent variable group having a sample probability larger than the population probability is over sampled, with the over sampled group having understated classification and prediction error rates". The population frequency has not exceeded 0.75% in the United States since 1934 according to statistics provided by Dun and Bradstreet (1982). As a consequence of this argument Zmijewski expects the following if the sample bias is included: "higher distressed firm sample frequency rates cause lower distressed firm error rates". From this we would expect samples reflecting the failure frequency rate to have higher rates in their prediction models. Table 2.1 shows results from Zmijewski (1984) for a probit model on different samples.

The results imply that if adjustments in the analytical techniques are made then a paired design may be appropriate. This involves assigning prior probabilities to group membership. The models developed on unbiased data samples will have higher misclassification rates but the should be more representative of the true classification accuracy of the model.

Table 2.1 Results of a study by Zmijewski (1984).

TABLE 4

Comparison of Unweighted and WESML Probit Model Classifications Across Alternative Estimation Samples[a]

	Choice Based Samples[b] 40 40 (.500)	40 100 (.286)	40 200 (.167)	40 400 (.091)	40 600 (.063)	40 800 (.048)	Pearson Correlation Coefficient[c]
Panel A — Unweighted Probit							
CLASSIFICATIONS[d]							
BANKRUPT	92.5	77.5	77.5	72.5	65.0	62.6	.949[h]
NONBANKRUPT	100.0	100.0	99.0	99.3	99.5	99.5	.681[j]
OVERALL	96.3	93.6	95.4	96.8	97.3	97.7	NA
CHI SQUARE[e]	65.17[h]	95.10[h]	158.10[h]	267.05[h]	345.88[h]	420.93[h]	NA
Panel B — WESML Probit							
CLASSIFICATIONS[d]							
BANKRUPT	52.5	47.5	50.0	48.0	42.5	42.5	.797[i]
NONBANKRUPT	100.0	100.0	99.5	100.0	99.8	99.9	.310
OVERALL	76.3	85.0	91.3	95.0	96.3	97.1	NA
CHI SQUARE[e]	25.83[h]	60.98[h]	96.19[h]	176.38[h]	230.62[h]	306.32[h]	NA
Panel C — Comparisons							
PERCENT THE SAME[f]	80.0	91.4	95.0	96.8	98.3	98.7	−.989[h]
CHI-SQUARE[g]	30.23[h]	121.6[h]	136.49[h]	225.17[h]	367.84[h]	485.23[h]	−.833[h]

[a] Classifications are based on the estimated probabilities of each estimation sample for a given technique using a .5 probability cutoff, that is, firms with probabilities greater than or equal to (less than) .5 are classified as bankrupt (nonbankrupt).
[b] Number of bankrupt number of nonbankrupt firms in the choice based estimation sample. The bankrupt firm sample frequency rate (number of bankrupt firms/total number of sample firms) is reported in parentheses.
[c] Pearson correlation coefficients between the estimation sample frequency rate and the result reported in the corresponding row. NA indicates that the correlation test is not applicable.
[d] Percentage of firms correctly classified.
[e] Chi square test comparing actual with classified states.
[f] The percentage of firms classified to be in the same state by both the *WESML* and unadjusted probit assessments.
[g] Chi square test comparing the *WESML* probit classifications with the unadjusted probit classifications.
[h] Significant at the .01 level.
[i] Significant at the .05 level.
[j] Significant at the .10 level.

Lau's model

Lau (1987) has presented a model which extends previous corporate failure prediction models in two ways:

1. instead of the conventional failure/ non-failure dichotomy, five financial states are used to approximate the continuum of corporate financial health, and
2. instead of classifying a firm into a certain financial state, the new model will estimate the probabilities that a firm will enter each of the five financial states.

The ranked probability scoring rule was used to evaluate the quality of such probabilistic predictions. The first extension enables the prediction of prefailure distress in addition to ultimate failure. The second extension conforms with more recent advances in prediction methodologies. The five financial states are:

state 0 = financial stability

state 1 = omitting or reducing dividend payments

state 2 = technical default and default on loan payments

state 3 = protection under Chapter X or XI of the Bankruptcy Act

state 4 = bankruptcy and liquidation.

States 1 to 4 were states of increasing severity of financial distress. The prediction models were constructed with an original sample and then tested with a holdout sample. Each sample contained 350 firms in the financially healthy state 0, and 20, 15, 10, and 5 firms in states 1, 2, 3, and 4 respectively. These firms were selected as follows.

1. State 0 firms. From the Compustal tapes, 350 firms which were financially healthy in 1976 (1977) were selected for the original (holdout) sample. Every firm met the following conditions: (i) its assets-size fell in the same range ($1.6 million to $120 million) as

that of the financially distress firms; (ii) its 1971-1977 financial reports were available; and (iii) it experienced no financial distress or financial loss between 1972 and 1977.

2. State 1 firms. The Compustat tapes were used to generate a list of firms that reduced their annual dividend rate per share in 1976 (for the original sample) or 1977 (for the holdout sample) by more than 40% below that of the previous year. From this list, 20 firms were selected for each sample.

3. State 2 firms. The Wall Street Journal Index (WSJI) and the Standard and Poor Stock Reports were used to compile a list of firms that had either filed for protection under Chapter X/XI during 1977 to 1980 or had C-rated bonds. The 10-K reports of each of these firms were examined to identify those that defaulted loan interest and/or principal payments during 1976 and 1977. Fifteen loan-defaulting firms were so obtained for each of the two sample.

4. State 3 and 4 firms. From the WSJI list of bankrupt and Chapter-X/XI firms, for each sample, 10 Chapter-X/XI and 5 bankrupt firms that had publicly available 10-K reports were selected. Lau used three groups of variable: financial flexibility variables (contained 7 variables), two trend variables and indicator of current financial state. These variables are summarized in Table 2.2 and explained below. Lau's financial distress prediction models were constructed using multinomial logit analysis . Considered the problem in which all firms will enter one of J = 5 states. Each firm's destiny is predicted by K = 10 explanatory variables, designated x_1, x_2,...., x_{10}. Defining P_j as the probability that a given firm will eventually enter state j, the logit model postulates that the P_j's of the firm can be estimated as follows:

(i) compute $Z_j = b_{j1}x_1 + b_{j2}x_2 + \ldots\ldots + b_{j,10}x_{10}$

for each state $j = 0$ to 4,

(ii) then $P_j = \exp(Z_j) \,/\, \sum_{j=1}^{J} \exp(Z_j)$.

The coefficient b_{jk} can be considered as the effect of the kth explanatory variable on a firm's probability of entering state j. Predictive models with three different predictive horizons were constructed. A "year 1" prediction model was constructed with 1974/1975 financial information to predict financial distress in 1976, and similar "year 2" and "year 3" models were constructed with 1973/1974 and 1972/1973 financial information respectively to predict financial distress in 1976. The holdout sample was used to test the ability of these models to predict 1977 distress. Lau used the QUAIL program by Berkman et al (1979) to construct three logit prediction models, one for each prediction horizon. Each model has five logit functions, one for predicting each of the five states. The expected sign of each coefficient in each logit function depends on the effect that a variable has on a firm's final state. Lau points out that applying a probabilistic prediction model to a group of n firms gives n probabilistic prediction scores, and the prediction performance is represented by the sum of these n scores (SS_n) as well as the ratio SS_n/n (since n is the maximum possible sum of scores). Lau's results are presented in Table 2.3, which gives the SS_n for each of the five groups of firms and for the entire set of 400 firms. For example, it indicates that the probabilistic predictions produced for 15 state-2 firms in the original sample by the year-1 multinomial logit analysis prediction model earned a total of 14.38 out of a maximum possible score of 15. Except for the state-4 firms in the holdout sample, Table 2.3 indicates that the score earned by each group of firms is close to the maximum possible score. For comparison, multiple discriminant analysis was applied to the same data set, and the SS_n's for the entire set of 400 firms are given in the last column of Table 2.3. It can be seen that multinomial logit analysis outperforms multiple discriminant analysis in every case, with larger differences in the holdout sample.

The results of a multinomial logit analysis were poor in comparison with those reported in earlier two-state models, but this is partly due to the overstatement of predictive accuracy of the earlier work, and also because a five state model demands more from the data and could itself be a reason for the poorer results.

Table 2.2 Summary of explanatory variables used in Lau's (1987) five-state financial distress prediction.

Variable Number	Brief Definition of Variable	Abbreviation	Nature of the Variable
Group 1: Financial Flexibility Variables			
X_1	= 1 if one of the firm's loan agreements contains 3 or more restrictive terms and the loan's interest is above the prime rate. = 0 otherwise.	*LRT*	dichotomous
X_2	(Firm's Debt-Equity Ratio) ÷ (Industry Debt-Equity Ratio).	*DER*	ratio
X_3	Working Capital Flow/Total Debt.	*WFTD*	ratio
X_4	Stock Price Trend = $\frac{(H_t - H_{t-1}) + (L_t - L_{t-1})}{H_t + H_{t-1} + L_t + L_{t-1}}$ where H_t and L_t are, respectively, the high and low values of the range of stock prices in year t.	*TCSP*	ratio
X_5	(Firm's Operating Expense to Sales Ratio) ÷ (Industry's Operating Expense to Sales Ratio).	*OPES*	ratio
X_6	= 1 if no dividend is being paid currently, = 0 otherwise.	*DCSD*	dichotomous
X_7	= 1 if the firm liquidates its operating assets in the period *and* there is no decreasing trend of earnings flow, = 0 otherwise.	*LOPA*	dichotomous
Group 2: Two Trend Variables			
X_8	Trend of Capital Expenditures = $\frac{(K_t - K_{t-1})}{(K_t + K_{t-1} + K_{t-2} + K_{t-3})/4}$ where K_t = capital expenditure in year t.	*TCEP*	ratio
X_9	Working-Capital Flow Trend = $\frac{(WF_t - WF_{t-1})}{(WF_t + WF_{t-1} + WF_{t-2} + WF_{t-3})/4}$ where WF_t = working capital in year t.	*TWF*	ratio
Group 3: Indicator of Current Financial State			
X_{10}	= 1 if dividend payments are omitted or reduced more than 40% in the period. = 0 otherwise.	*DVD*	dichotomous

Table 2.3 Aggregate probabilistic prediction scores earned for different groups of firms for Lau's model (1987).

Prediction Horizon	Multinomial Logit Models: Firms in State[1] 0 ($n = 350$)	1 ($n = 20$)	2 ($n = 15$)	3 ($n = 10$)	4 ($n = 5$)	All Firms ($n = 400$)	Multiple Discriminant Models All Firms ($n = 400$)
Original Sample[2]							
Year-1 Model	349.2	18.67	14.38	9.45	4.64	396.3	391.3
Year-2 Model	347.7	17.12	13.21	7.47	5.00	390.5	385.8
Year-3 Model	347.1	16.48	12.53	7.69	4.34	388.1	379.6
Holdout Sample[3]							
Year-1 Model	336.0	17.33	13.25	7.80	1.74	376.1	332.1
Year-2 Model	335.2	16.87	11.30	7.82	2.52	373.7	369.1
Year-3 Model	334.9	16.52	12.35	7.52	2.84	374.2	365.7

[1] State definitions—five-state financial distress models:
State 0: financial stability;
State 1: omitting or reducing dividend payments;
State 2: technical default and default on loan payments;
State 3: protection under Chapter X or XI of the Bankruptcy Act;
State 4: bankruptcy and liquidation.

[2] Original sample: financial information from 1972-75 is used to predict firms in financial distress in 1976.

[3] Holdout sample: financial information from 1973-76 is used to predict firms in financial distress in 1977.

Keasey and Watson model

Keasey and Watson (1987) , used logit function to determined whether a model utilising a number of nonfinancial variables, either alone, or in conjunction with financial ratios, was able to predict small company failure more accurately than models based upon financial ratios only. For the logit functions the dependent variable was failure/non-failure and the set of indepent variables were as follows:

model 1 = financial ratios only

model 2 = nonfinancial information only

model 3 = financial ratios and nonfinancial information.

The financial ratios used in models 1 and 3 consist of 28 ratios, covering various aspects of company performance such as profitability, liquidity and gearing. The non-financial variables included are number of directors, time lag in submitting accounts to Companies House, audit qualifications and the presence of a secured loan. The sample of 146 companies (73 failures and 73 non-failures) used to obtain the univariate results was utilised in their study to obtain the initial logit functions. Information on a further 20 companies (10 failures and 10 non-failures) was obtained for use in holdout tests. No attempt to incorporate the relative costs of misclassification of failed and non-failed companies was undertaken. The financial and non-financial information for failed companies has been taken from the last three years of published accounts available before failure, therefore not restricted to a common period prior to failure. From a practical decision-making viewpoint this procedure of basing the logit functions upon the most recent information that is available for each company seems sensible for two reasons. First, the practical decision-maker cannot exclude companies merely because they have not submitted their latest set of accounts to Companies House. Second, it recognizes that a practical decision model can only utilize information which is available prior to failure.

The correct classification results were 76.7% for model 1, 75.3% for model 2 and 82.2% for model 3. So comparing the results of model 1 and 2 it is apparent that the non-financial information contained in model 2 does not succeed in correctly classifying a greater number of cases than the benchmark model 1. However, its poorer performance in terms of classificatory success for the original sample is marginal. The overall correct classification rates for the holdout sample shows that, the non-financial data-model 2, provides a better overall prediction rate (65%). Furthermore, the more extensive model 3 does not appear to provide a better overall prediction rate (65%) than model 2. They conclude that marginally better predictions concerning small company failure can be achieved by the use of these non-financial variables.

Barnes's model

Barnes (1990), used multivariate discriminant analysis to predict takeovers. Barnes points to three factors effecting predictive ability. These are: (i) the strict statistical assumptions on which the estimating procedures are based, (ii) further statistical implications arising from the way in which the sample is chosen, and (iii) the predictive application of the model which includes, particularly, its stability over time. Data concerning 92 takeover bids of UK quoted companies during the years 1986-1987 were obtained (mergers announced prior to the October 1987 crash). Each company was matched with a non-acquired listed company within the same industrial sector whose market capitalisation immediately prior to the merger was the nearest. Nine basic financial ratios for each company two years prior to the merger were obtained. However, the ratios themselves were not used in the discriminant model. Instead, the ratio between it and the relevant sector average, the industry-relative ratio (which is defined as the ratio of a firm's financial ratio relative to the mean value for that ratio in the firm's industry at a point in time) was used. Barnes also used factor analysis, in order to eliminate the effects of statistical multicolinearity and the overlapping nature of some of the nine ratios. Five factors were found to explain 91.48 percent of the variance in the original data matrix. His model predicted 68.48 percent correctly. The predictive accuracy of Barnes' model was tested on a further group of 37 acquired companies and 37 matched non-acquired companies. Here the model predicted 74.3 percent correctly. On UK data using multivariate discriminant analysis, he achieves good predictive ability but does not test his model on a subsequent period due to the stock market crash in October 1987.

2.5 Summary and Implications

In this chapter the concept of failure was introduced and a number of failure prediction studies have been discussed, including the seminal studies of univariate analysis (Beaver, 1966) and multivariate analysis (Altman, 1968). An overview of subsequent research has also been given, predominantly using multivariate discriminant models (MDA).

An examination of the methodology used in the earlier bankruptcy prediction studies shows that there are three principal methodological flaws which make the reported prediction accuracies unreliable.

pitfalls in MDA With the exception of Ohlson (1980), Zmijewski (1984), Lau (1987) and Keasey and Watson (1987), the multivariate financial prediction studies reviewed here have used MDA for model construction. Eisenbeis (1977) provided a detailed discussion of several flaws in the way MDA has been used, and of the limitations which result. These are as follows :

(i) Distribution assumption of the variables : The MDA technique assumes that the explanatory variables are multivariate normally distributed, and Lachenbruch (1975) has shown that both linear and quadratic discriminant analysis are quite sensitive to this assumption. However, most of the MDA prediction studies ignored the need to test for the multivariate normality of their explanatory variables.

(ii) Choice of a priori probabilities : The importance of assigning correct a priori probabilities to the various discriminant groups was overlooked in earlier studies. Most researchers simply assumed that group membership is equally likely among possible groups, even though in the actual population the number of surviving companies is usually much higher than the number of non-surviving companies.

(iii) Interpretation of the significance of explanatory variables : The earlier financial prediction models using MDA either overlooked the interpretation of the significance of the individual variables or have interpreted it incorrectly. Consider a discriminant function of the form $Z = a_0 + \sum_{i=1}^{k} a_i X_i$, where the X_i's are the explanatory variables. An extension from multiple regression suggests that the importance of an X_i is indicated by its standardised coefficient $a_i X_i$. However, Eisenbeis (1977) pointed out that, unlike the coefficients in multiple regression, "the discriminant function coefficients are not unique; only their ratios are. Therefore, it is not possible, nor does it make any sense to test, as in the case with regression analysis, whether a particular discriminant function coefficient is equal to zero or any other value". Eisenbeis then reviewed and evaluated several methods that have been proposed in the literature to determine the relative importance of the individual explanatory variables.

(iv) <u>Assessment of classification errors</u> : In most of the early studies, the hold-out sample used for cross validation was drawn from the same period as the analysis sample used to derive the discriminant function, and the cross validation test was then presented as a prediction test. In fact, a validation test using a hold-out sample from the same test period merely validates ex-post discrimination. It does not validate the model's ability to predict for future periods. Therefore, those studies that presented the cross validation tests as prediction test may have over-estimated the predictive ability of their models.

Ohlson (1980) and Keasey and Watson (1987) applied a new technique to multivariate bankruptcy modelling by estimating logistic models. Lau (1987) uses multinomial logit. Each expected that a logistic and multinomial logit models would improve the results since the data provide a better fit for the assumptions of the technique. The results of neither study bear this out. However, the strong significance of the estimates for these models, the pattern of significance of the financial attributes and the information content transformed logit probabilities as a financial risk measure appear to be the main contributions of these techniques.

Sampling The use of matched samples in the majority of previous studies will have resulted in sample selection bias in the absence of a suitable estimation procedure. Therefore, earlier studies may be said to suffer in this respect from inconsistent and biased estimates of model predictions.

Definition of 'failure' Many previous prediction models have defined failure narrowly as bankruptcy. This narrow definition of failure leads to the restricted population sizes used in many previous studies.

Implications

The methodological critiques of Eisenbeis (1977) and Palepu (1986) have identified a number of shortcomings in failure prediction models and particularly with regard to the use of discriminant techniques, sampling and the definition of failure. In this thesis, particular attention is paid to the improvements of statistical method, especially with respect to the three aspects mentioned above that is :

- the extension of MDA and logit analysis to survival modelling in the context of censored observations,
- the use of unbalanced groups of survivors and non-survivors, leading to a randomly stratified sampling technique,
- the introduction of a likelihood solution to sample construction when there are companies which cease trading for reasons other than technical bankruptcy, such as companies which are taken over.

These three issues are discussed in the remainder of this thesis.

CHAPTER THREE

THE DATA SET AND EXPLORATORY ANALYSIS OF THE FINANCIAL RATIOS

3.1 The Sample

The analysis reported in this thesis is based on the available set of non-surviving and surviving companies covered by EXSTAT - i.e. all UK companies of interest to the investing institutions (see appendix 1 for more information about EXSTAT). Previous studies of corporate failure have used smaller size data bases of companies that were often a mix of manufacturing, merchandising, and various other industries (Beaver, 1966 studied 158 companies, thirty being nonmanfacturing companies, e.g. twelve merchandising companies and various other types of company). In constructing the sample for this study, two general guidelines were followed. First, a paired-sample technique is not employed - each non-surviving company is not matched in the analysis with a surviving company. Second, only industrial companies are considered in this study.

The information covers the period 1971 to 1984. As mentioned above, the companies were selected from EXSTAT's industrial sectors (codes 19 to 34). Table 3.1 shows the sectors in which the surviving and non-surviving group of companies were operating.

Table 3.1 Industrial classification.

Name of sector	Sector group No.
Electricals (excluding radio and T.V.)	19
Cold formed fastenings	20
Founders and stampers	21
Industrial plant, Engines and compressors	22
Mechanical Handling	23
Pumps and values	24
Steel and chemical plant	25
Wires and Ropes	26
Misc. Mechanical Engineering	27
Machine and other Tools	28
Misc. Engineering contractors	29
Heating and Ventilating	30
Instruments	31
Metallurgy	32
Special steels	33
Misc. metal forming	34

The Non-surviving Companies

Depending on one's definition of failure, various interpretations are possible. Reference to Table 3.2 (below) shows that if failure is taken solely as bankruptcy, then only 21 companies (plus one which did not survive for one year and, therefore, did not publish more than one set of financial statements) did not survive. However if a broader definition is used, then various other categories might be included (Martin, 1975) which means that up to 104 companies could be considered as non-survivors.

Table 3.2 Classification of 104 non-surviving companies

Classification		Number of sample
Bankruptcy		22
Other liquidations	19	
Mergers or takeovers	63	
		82
		104

An analysis of the data for non-survivors is given in Table 3.3. This shows for each of the subclasses - "bankrupt", "merged" and "others" - (i) the number of companies failing in each of the years from 1971 to 1984, and (ii) the number of companies for which data was available in each year.

In addition, Table 3.4 shows the length of the time series available for analysis. It can be seen that a number of companies do not survive for five years. Previous researchers have tended to eliminate such companies from their analysis. That approach has not been followed here, although companies which did not survive for even one year were excluded as there would have been no data on which to base prediction. This is an important point in the context of survival analysis (see Chapter 6) and also affects sample selection in the application of discriminant analysis and logit analysis (see Chapters 4 and 5). Consequently, the analysis reported in the thesis omits 1 bankrupt company and 8 merged companies. The sample sizes used in analysis were as follows :

Bankruptcy	21
Other liquidations	19
Mergers or takeovers	55
	95

Appendix 2 gives a listing by name of non-surviving companies.

Table 3.3 Number of non-surviving companies for which data was available.

Year	Bankrupt		Merged		Other	
	i	ii	i	ii	i	ii
1971	0	1	0	7	0	2
1972	0	4	0	15	0	6
1973	0	3	0	13	0	5
1974	0	4	0	16	0	6
1975	0	5	0	23	0	10
1976	1	8	6	30	0	14
1977	1	15	10	45	1	16
1978	5	21	5	39	0	16
1979	1	16	4	35	0	18
1980	1	14	11	31	1	18
1981	2	13	9	24	5	17
1982	4	11	4	14	4	12
1983	4	7	9	14	2	8
1984	3	3	5	5	6	6
Total	22	—	63	—	19	—

Key

(i) the number of companies failing in each of the years.

(ii) the number of companies for which data was available in each year.

Table 3.4 Length of time series for non-surviving companies

No. of reporting periods	Bankrupt	Merged	Other	Total
1	1	8	0	9
2	3	7	0	10
3	1	8	0	9
4	0	6	0	6
5	3	8	1	12
6	6	10	5	21
7	3	5	5	13
8	3	3	3	9
9	1	1	1	3
10	1	5	0	6
11		1	0	1
12		1	2	3
13			1	1
14			1	1
Total	22	63	19	104

The Surviving Companies

The sample of continuing or surviving companies was constructed in a way whereby there was no matching with non-survivors by size or financial year, nor was the number of companies equal in the two groups. Taffler (1982) argued that restricting the size of the surviving sample to that of the non-surviving set only served to restrict the total sample size and degrees of freedom. He argued that the statistical methods only require separate multivariate normality in the constituent groups together with equality of their variance-covariance matrices. Therefore there is no need for the surviving sample to be exactly the same size as the non-surviving set.

There were 359 surviving companies in the sectors previously indicated at the date when the EXSTAT tape was compiled. Earlier, in Section 2.4.1.2, we mentioned that Taffler (1982) also explicitly recognized that a continuing company is not necessarily financially healthy and so, in his study, the surviving sample was made up of healthy solvent companies. Taffler used a group of investment analysts of a leading company of London stockbrokers to judge whether a continuing company is fully solvent or not. However, in our study, this step of selecting only healthy solvent companies was ignored as it would necessitate external assistance not available during the selection process.

However, amongst the surviving companies were some which either had been existence for only one year, or which had not reported in 1984, the last year of data on the EXSTAT tape used. Table 3.5 shows the availability of data for such companies in each year, and the length of the time series available for analysis. As with the non-survivors, a number of survivors had been in existence at the date of censoring for less than 5 years. As mentioned before, the companies with short lives were not excluded, except for nine which reported only in the last year. For these, it was not possible to include even one lagged observation in the analysis. In addition, thirteen companies which were coded as survivors had not reported in either 1983 or 1984. These were excluded from the analysis under stratified sampling (see Chapter 4). Consequently the differing samples used in this thesis are based on the following :

Survivors	337
Survivors not reporting in 1983/1984	13
	350
Survivors reporting in 1984 only	9
	359

An issue which has been considered by others (Barron, 1986) concerns the heterogeneity of reporting dates. That is, companies may change their reporting date and reporting period, i.e. for the year ended 30 / 6 / 81, for example, to a subsequent period of 9 months to 31 / 3 / 82 or, perhaps, 15 months to 30 / 9 / 82. Although this was taken in to account in computing financial ratios by annualising the ratio, this affects the structure of the data which is assumed for the majority of companies to be one of yearly reporting. In this thesis, it is assumed that the time series follows annual intervals in spite of the above. This is justified by the following analysis : The number of reporting periods was compared to the length of the time series in years. There was a variation in only 19 out of 359 cases. In 8 cases, there was one more set of reported accounts than years of existence. In 11 cases, there was one less. In other words, there were no substantial timing effects *in the longer-term.* In view of the fact that one company, for example, reported twice in 1972, not in 1975, twice in 1977, not in 1981, and twice in 1984, the support for the assumption of regular reporting in the longer-term allows us to deal effectively with a problem that, nevertheless, has shorter-term implications.

Appendix 3 gives a listing by name of surviving companies.

Table 3.5 Number of surviving companies for which data was available and the length of time series.

No. of surviving companies in each year	
Year	Number of companies
1971	20
1972	49
1973	58
1974	64
1975	83
1976	153
1977	275
1978	317
1979	320
1980	309
1981	319
1982	333
1983	350
1984	346

Length of time series	
No. of reporting periods	Number of companies
1	9
2	13
3	10
4	4
5	3
6	6
7	50
8	119
9	60
10	22
11	10
12	14
13	18
14	21
Total	359

3.2 Collection of Financial Ratio Data

In this study as in many previous studies the independent variables are various financial ratios which are calculated from financial data disclosed by the companies under study. No attempt was made here to develop any new ratios. The financial ratios used here are ones that were regarded as notable by a collection of current accounting and finance texts. The emphasis was not on trying to develop new financial concepts but to evaluate carefully the existing concepts. The variables included in the study were chosen on the basis of their

(1) popularity in the literature,

(2) potential relevancy to the study (Altman, 1968).

The 23 financial ratios which were decided on for this study are listed in Table 3.6.

It should be noted that financial ratios are constructed from accounting information disclosed by companies, and used as indicators of financial structure or performance. Generally, the underlying accounting information relates to residual balances at a particular point in time (such as the amount of liquid funds held by a company at the close of business on, say, 31st December) or to transactions accumulated over a period of time (e.g. the salaries and wages paid by the company, or the profits calculated by the company, for the year from 1st January to 31st December). Companies disclose such information from one year to the next, generally for the same period and at the same closing date. Hence, financial ratios can be viewed as repeated measures. However, there are a number of issues to be considered in this respect:

Table 3.6 List of financial ratios

Financial Ratio	
1. Net income to sales	(NI/S)
2. Funds flow to net worth	(FF/NW)
3. Funds flow to total assets	(FF/TA)
4. Net income to total assets	(NI/TA)
5. Net income to net worth	(NI/NW)
6. EBIT (earnings before interest & tax) to sales	(EBIT/S)
7. EBIT (earnings before interest & tax) to total assets	(EBIT/TA)
8. Quick assets to total assets	(QA/TA)
9. Funds flow to sales	(FF/S)
10. Current assets to total assets	(CA/TA)
11. Net worth to sales	(NW/S)
12. Sales to total assets	(S/TA)
13. Total assets to net worth	(TA/NW)
14. Funds flow to current liabilities	(FF/C.LIB)
15. Retained earnings to total assets	(RE/TA)
16. Current assets to current liabilities	(CA/C.LIB)
17. Quick assets to current liabilities	(QA/C.LIB)
18. Current liabilities to net worth	(C.LIB/NW)
19. Current liabilities to total assets	(C.LIB/TA)
20. Cash to sales	(CASH/S)
21. Cash to total assets	(CASH/TA)
22. Current assets to sales	(CA/S)
23. Quick assets to sales	(QA/S)

(i) As mentioned in the Section 3.1, occasionally, a company will alter its reporting date and disclose information for 9 months or 15 months, rather than the usual calendar year; thus, these repeated measures are characterized by some irregularity.

(ii) Although, the greater proportion of companies report for the calendar year to 31st December, there is heterogeneity in reporting dates in countries such as the U.K.. In a sense, the sample revolves through a one year cycle, with all "current" observations being updated during that period. In the context of time series analysis, this is an important issue as "mean" effects may be estimated on each occasion that a single company releases new information whilst, at any one point in time, the latest information relating to the sample will cover periods beginning up to two years beforehand.

3.3 Mean Effects and the Influence of Censoring

In this Section, we provide some exploratory analysis of the general time series behavior of ratios, in the light of the features discussed in Section 3.2. Generally, we might consider an observation for the ith company at time t to be a linear combination of past terms

$$Y_{i,t} = \beta_1 Y_{i,t-1} + \beta_2 Y_{i,t-2} \cdots\cdots$$

Allowing for an effect that is attributable to conditions influencing all companies, we rewrite this as

$$Y_{i,t} = \bar{Y}_t + \bar{Y}_{t-1} + \beta_1(Y_{i,t-1} - \bar{Y}_{t-1})\ldots\ldots\ldots$$

where

$\bar{Y}_{t-1}$ is a mean effect (the mean of observation for all companies at t-1)

$\beta_1(Y_{i,t-1} - \bar{Y}_{t-1})$ is a company effect relating to company i at t-1 and

$\bar{Y}_t$ is a mean effect for the current period which, for example, could be estimated from observations on other companies for the current period.

This is a generalised representation of the view that current outcomes are explained by events influencing all companies during the current period and past periods, modelled by some kind of transfer function, plus an effect attributable to the company's own past (in this case, we assume a systematic company effect, but this could well be simplified to a random term).

However, our preliminary analysis showed that such mean effect tend to be far less influential than expected, particularly in comparison with the influence of censoring. We estimated mean effect for each month, in order to start to overcome problems of non-synchronous reporting and changes in reporting date. Then, we generated monthly estimates for each company of the ratio $Y = X_1/X_2$ by assuming that X_1 and X_2 are each described by straight lines between reporting dates (not a very adventurous approach, but a straightforward starting point) by using the following straight line equation

$$\frac{Y_2 - Y_1}{X_2 - X_1} = \frac{Y - Y_1}{X - X_1}$$

The mean was estimated as the median monthly observation for the sample. As inferred above, the kind of perfect foresight assumed here is relatively simple, and more interesting approaches could be adopted to take account of information in past time series, the incremental evolution of the sample data set, and so on. Nevertheless, it is interesting that the median follows a path in the longer term that is well-described by the effects of censoring. For example, for the median Current Assets/Current Liabilities (CA/C.LIB) ratio for censored companies (i.e. survivors), we note a relatively stable median seemingly uninfluenced by temporal conditions in the economy for most of the period, but declining towards the end (see Figure 3.1 below). This decline in characteristic of censored data, due to the inclusion of failing (but not yet failed) companies in the censored sample. This can also be seen by plotting CA/C.LIB against termination time (i.e. the number of periods before failure). The ratio declines as failure approaches (see Figure 3.2 below). When we consider the mean effect for failed (and some taken-over companies), the proportion of observations relating to companies which are close to failure increases as the date of censoring approaches and, accordingly, the mean falls. For the median Net Income / Sales (NI/S) and Net Income / Total Assets (NI/TA) ratios for censored companies the results shown in Figures 3.3 and 3.4 for the NI/S ratio, 3.5 and 3.6 for the NI/TA ratio indicated properties similar to the CA/C.LIB ratio. Therefore it is reasonable to arrive at the same conclusion for these ratios as for the CA/C.LIB ratio.

Figure 3.1 The Current Assets / Current Liabilities ratios

for surviving companies and the median ratio

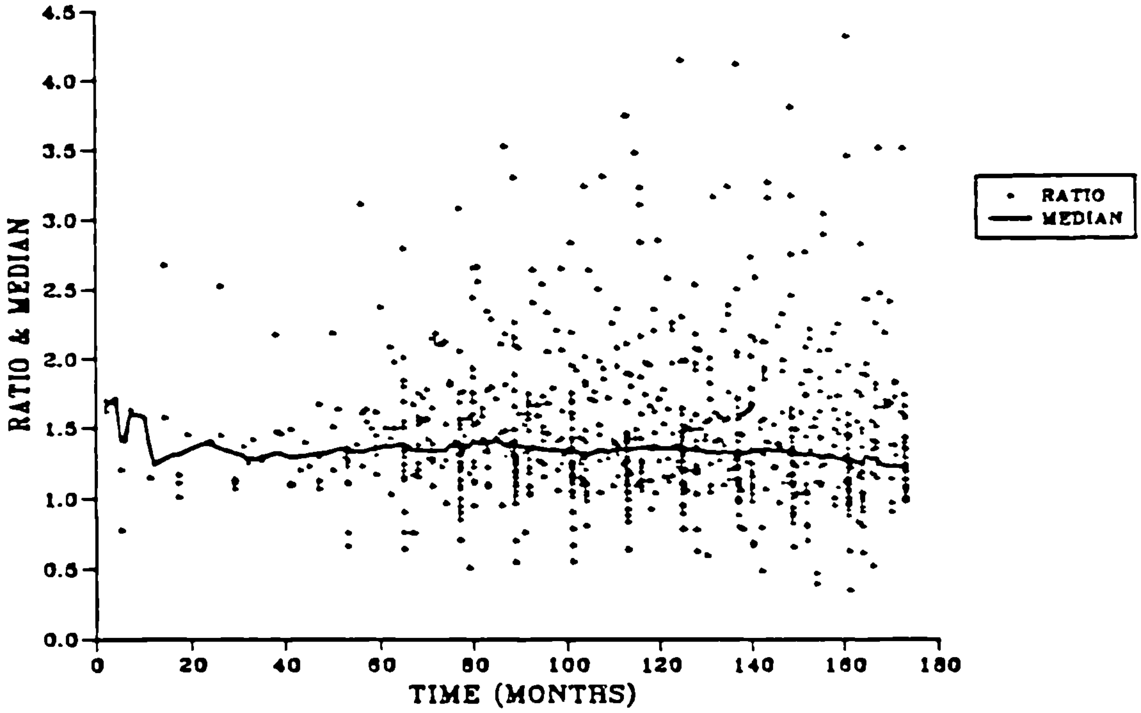


Figure 3.2 The Current Assets / Current Liabilities ratio vs.

Termination time for surviving companies

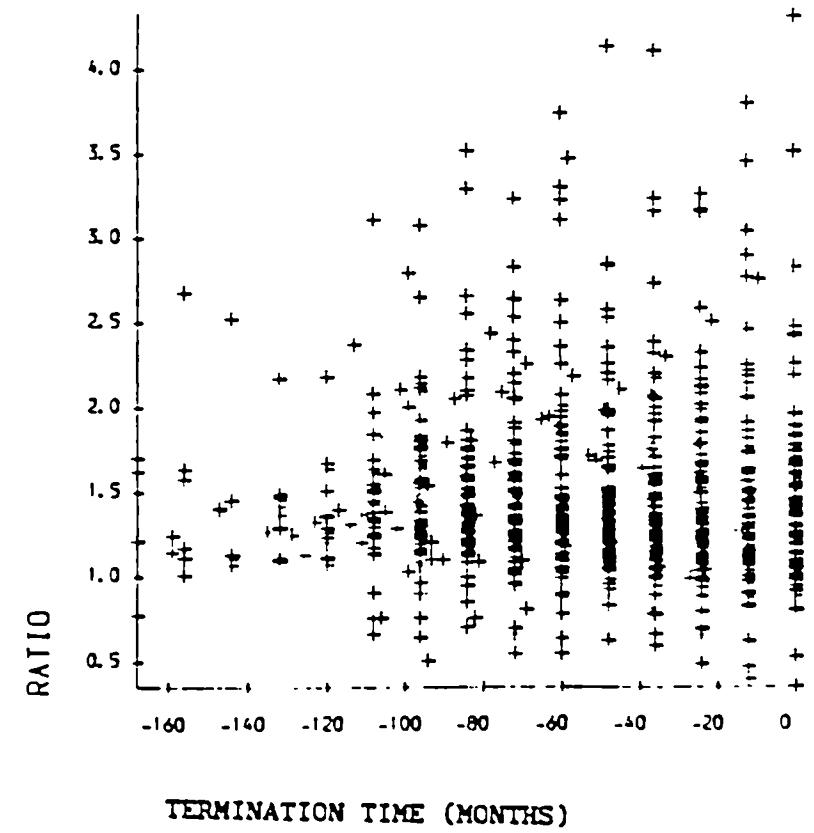


Figure 3.3 The Net Income/Sales ratios for surviving companies and the median ratio.

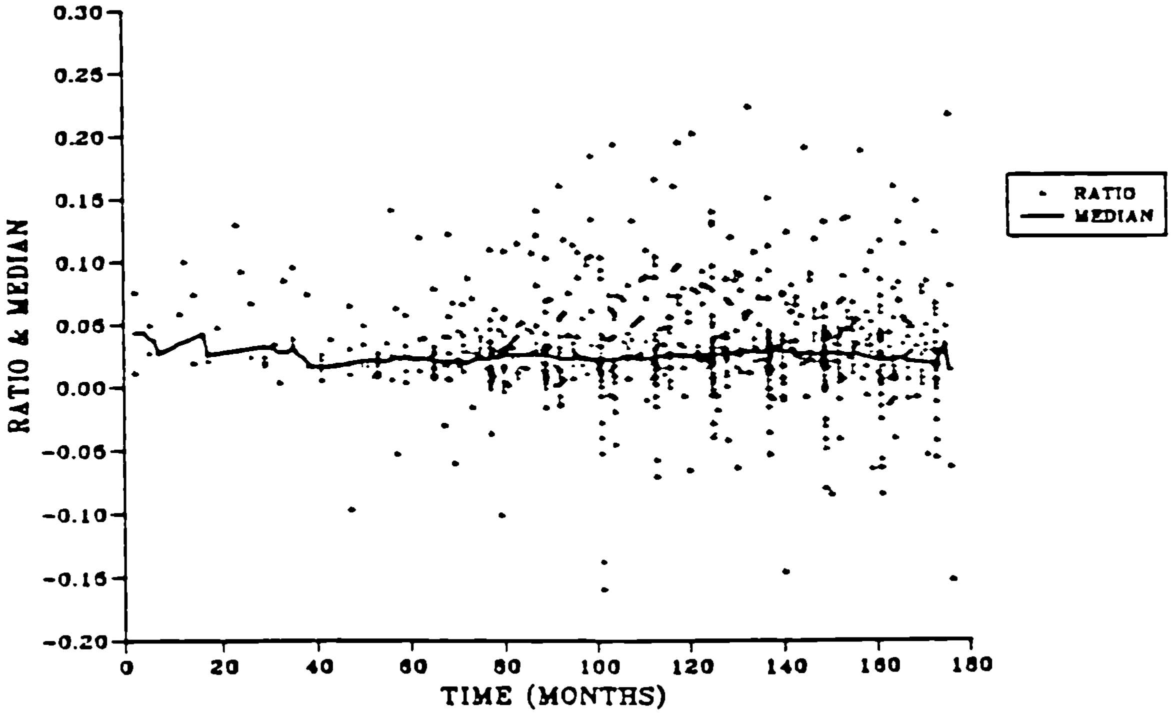


Figure 3.4 The Net Income/Sales ratio vs. Termination time for surviving companies

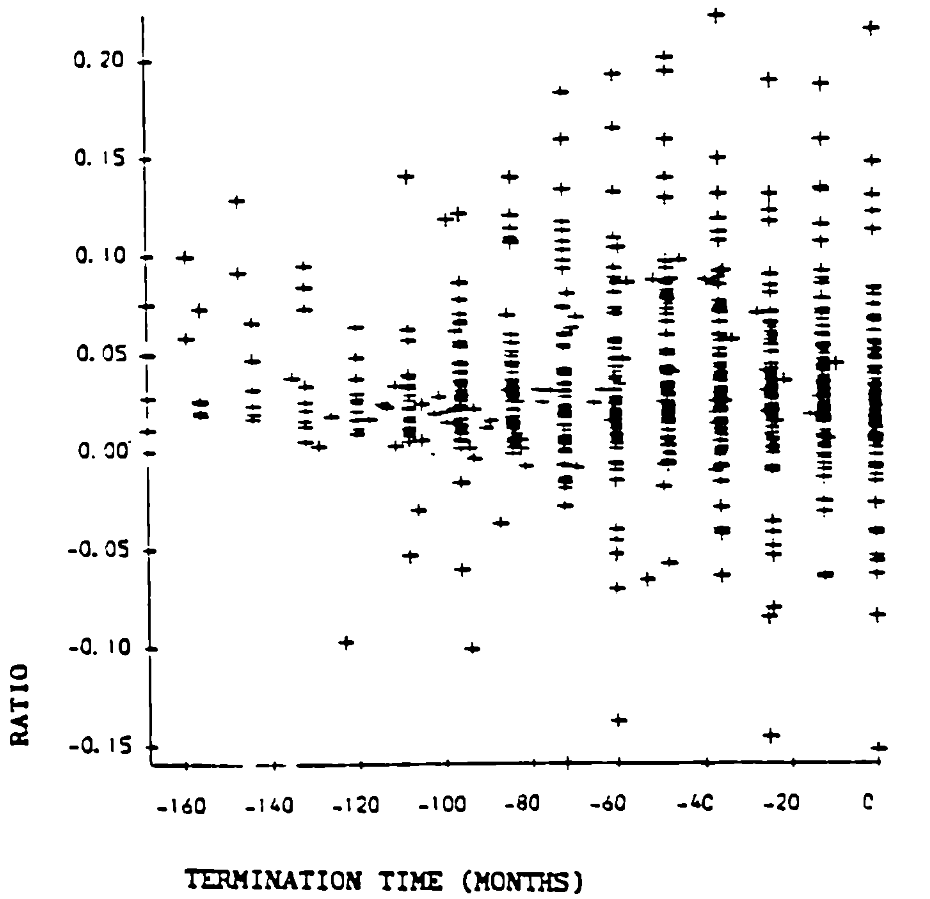


Figure 3.5 The Net Income/Total Assets ratios for surviving companies and the median ratio

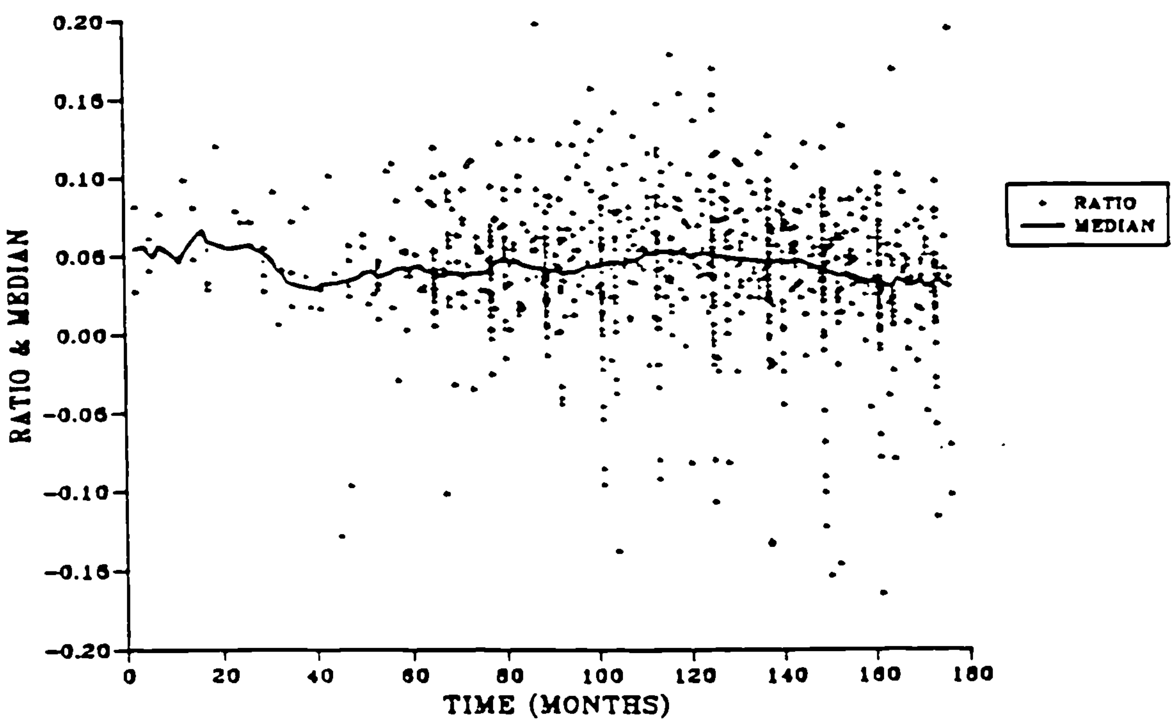
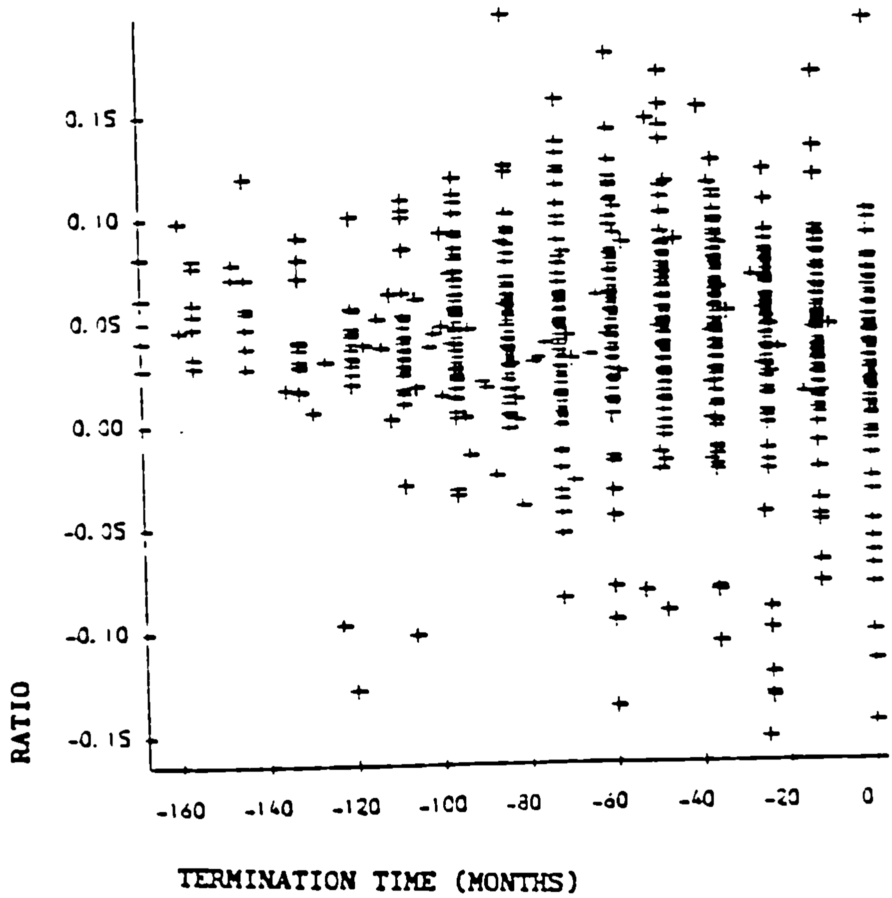


Figure 3.6 The Net Income/Total Assets ratio vs. Termination time for surviving companies



3.4 Distribution of Financial Ratios

Much research using corporate financial indicators assumes that the data are normally distributed. There has been some empirical refutation of this assumption. In this Section, the normality assumption is subjected to extensive testing, and the statistical characteristics of the 23 financial ratios are investigated.

This Section reports an empirical examination of the underlying distribution of selected U.K. industrial company financial ratios. Industrial data are usually assumed to have a normal distribution in industrial company research and financial analysis. Most studies involving the analysis of financial ratios either implicitly or explicitly assume that results from the sample are statistically related to the corresponding population ratios on the basis of the central limit theorem. This theorem states, in brief, that sampling distributions approach normality as sample size increases. This property is the basis for researchers' normality assumption. Populations that are normally distributed have sampling distributions that follow the normal distribution. As a population diverges from normality, for this population, how large must n be so that the normal approximation is accurate enough? This Section examines a more basic threshold question: are the population distributions of the ratios used for analysis in this thesis normal?

Deakin (1976) tested for normality (and for transformation to normality) using the financial ratios of U.S. manufacturing companies for the 19 years from 1955 to 1973. He concludes that an assumption of normality cannot be supported from his research on industrial companies. Bougen and Drury (1980) investigated financial ratio distributions, using the financial ratios of U.K. companies from 45 different industries, for 1975, again concluding that an assumption of normality was found untenable for each ratio. Bedingfield et al (1985) observed that many financial ratios are skewed and non-normally distributed. Barnes (1982) recommended that when the basic assumption of ratio analysis , i.e. proportionality, is violated, non-normality will be found. precisely, he demonstrated that when the intercept of ratio Y/X is not equal to zero, the distribution of this ratio will be skewed.

For each ratio, annual measures are employed for central tendency, dispersion, skewness and kurtosis. Each measure is a partial description of the underlying distribution and provides an indication of its shape and form. An effective means of determining whether an empirical distribution follows an hypothesized and theoretical distribution is to compare their characteristics. The appropriate analytical technique is to apply the Kolmogorov-Smirnov or chi-square goodness of fit test or Shapiro-Wilk test. To test the normality assumption, a Kolmogorov-Smirnov (K-S) statistic (Conover, 1971) is employed here. Table 3.7 presents the analysis of the financial ratios for U.K. industrial companies for survivors and non-survivors combined. The standard deviation is presented as one measure of dispersion and is supplemented by the dimensionless coefficient of variation. The third moment about the mean measures skewness, or the symmetry of a distribution about its mean. If the mean and median of an empirical distribution diverge, the distribution is skewed. The sign of the skewness statistic indicates whether the distribution is positively or negatively skewed. Kurtosis is measured by the fourth moment about the mean to describe the peakedness or flatness of the distribution. The larger the index, the greater is the peakedness; the smaller the index, the flatter is the distribution's shape.

Table 3.7 Descriptive of distribution of financial ratios for survivors and non-survivors combined.

Financial ratio	Mean	Standard deivation	Coefficient of variation	Skewness	Kurtosis	K-S statis-tic
NI/S	0.03	0.14	4.41	-43.48	2308.38	16.40
FF/NW	0.23	0.18	0.82	-2.65	55.78	5.70
FF/TA	0.11	0.09	0.79	-4.43	64.37	5.22
NI/TA	0.04	0.07	1.60	-6.91	118.38	8.56
NI/NW	0.08	0.16	1.89	-6.47	143.26	10.44
EBIT/S	0.06	0.16	2.67	-33.36	1652.90	13.27
EBIT/TA	0.08	0.09	1.22	-3.76	55.14	5.11
QA/TA	0.37	0.12	0.32	1.09	3.20	3.62
FF/S	0.08	0.15	1.80	-38.55	1942.80	13.86
CA/TA	0.69	0.12	0.18	-0.53	0.95	2.63
NW/S	0.44	0.31	0.71	1.82	51.51	8.58
S/TA	1.38	0.70	0.50	11.10	210.71	10.09
TA/NW	2.16	1.02	0.47	3.20	24.30	10.69
FF/CL	0.32	0.29	0.92	-5.17	140.50	5.17
RE/TA	0.28	0.21	0.74	-5.40	81.42	5.02
CA/CL	1.91	0.78	0.41	3.25	23.62	7.54
QA/CL	1.02	0.54	0.53	4.01	31.07	9.52
CL/NW	0.95	0.83	0.87	3.45	20.63	10.66
CL/TA	0.40	0.15	0.38	4.11	72.48	3.44
CASH/S	0.05	0.12	2.40	11.23	225.42	20.14
CASH/TA	0.06	0.09	1.56	3.05	13.05	15.51
CA/S	0.55	0.25	0.45	10.45	229.16	8.99
QA/S	0.30	0.20	0.67	15.62	457.48	13.24

The results in Table 3.7 show that the skewness estimates for all ratios are significant and also indicate that there is a decided and rather extreme skew to the distribution of financial ratios. The kurtosis shows profound peakness of all ratios which have kurtosis significantly larger than the value for the normal distribution. There is a peakedness that in some cases reaches exaggerated proportions. Note that the approximate sample variance for skewness and kurtosis statistics are given by 6/N and 24/N respectively for a normal distribution (Snedecor and Cochran, 1980). Based on the skewness and kurtosis statistic the normal distribution seems to be a poor distribution for describing financial ratios. The Kolmogorov-Smirnov (K-S) statistics in column 7 of Table 3.7 supports this argument. The Kolmogorov-Smirnov (K-S) statistics indicated that the ratios are not normally distributed.

Deakin (1976) reports that, at times the square root and natural log of ratios are normally distributed even though the raw data may not be. These same two transformations are made on the financial ratios used in this study and separate Kolmogorov-Smirnov tests are repeated as we can see in Table 3.8. An examination of these additional data indicates that neither the square root nor the log transformation assures normality.

The sample moments in Tables 3.9 and 3.10 provide some information about the distribution of the financial ratios for non-survivors and survivors respectively. The results in Tables 3.9 and 3.10 indicates that there is an extreme skew to the distribution of financial ratios in both cases. The financial ratios examined have a nonsymmetric distribution. Given the extreme skewness and kurtosis measures, the null hypothesis of normality is rejected. The Kolmogorov-Smirnov (K-S) statistics in Tables 3.9 and 3.10 supports this argument. This is an issue which has been investigated for the first time in a recent article by Hopwood, Mckeown and Mutchler (1988) and, although the analysis in this thesis does not attempt to use innovative models allowing for non-normality as the major focus is with other methodological issues.

Table 3.8 K-S statistic test for the natural log and the square root of the ratios for survivors and non-survivors.

Financial ratio	K-S statistic for log	K-S statistic for square root
NI/S	23.17	19.45
FF/NW	8.20	6.59
FF/TA	7.57	6.19
NI/TA	10.75	9.54
NI/NW	15.98	12.15
EBIT/S	21.43	16.87
EBIT/TA	7.59	6.12
QA/TA	2.96	2.39
FF/S	22.18	17.83
CA/TA	5.42	3.95
NW/S	12.55	8.79
S/TA	4.51	5.66
TA/NW	9.35	9.79
FF/CL	11.02	6.98
RE/TA	9.98	6.83
CA/CL	2.68	4.91
QA/CL	3.65	6.14
CL/NW	8.57	9.52
CL/TA	2.97	2.14
CASH/S	17.37	8.82
CASH/TA	14.30	7.10
CA/S	3.11	5.11
QA/S	5.29	8.32

Table 3.9 Descriptive of distribution of financial ratios for non-survivors.

Financial ratio	Mean	Standard deivation	Coefficient of variation	Skewness	Kurtosis	K-S statistic
NI/S	0.04	0.08	2.20	3.82	30.41	4.58
FF/NW	0.19	0.25	1.31	-7.10	100.44	4.00
FF/TA	0.01	0.08	0.78	-0.98	4.31	1.98
NI/TA	0.04	0.06	1.54	-0.96	3.37	2.64
NI/NW	0.05	0.07	1.45	-10.44	157.53	5.82
EBIT/S	0.07	0.13	2.02	5.78	47.99	5.28
EBIT/TA	0.07	0.08	1.22	-0.13	1.73	1.29
QA/TA	0.34	0.11	0.32	0.77	2.48	1.54
FF/S	0.08	0.08	0.95	-2.68	17.52	2.80
CA/TA	0.68	0.13	0.19	-0.88	1.35	1.62
NW/S	0.50	0.42	0.85	4.70	27.49	5.71
S/TA	1.24	0.37	0.30	0.26	1.35	1.23
TA/NW	2.16	1.03	0.47	5.92	59.29	4.54
FF/CL	0.27	0.25	0.93	-0.17	6.09	2.08
RE/TA	0.26	0.17	0.67	0.23	1.54	1.94
CA/CL	1.83	0.94	0.52	5.82	52.15	3.97
QA/CL	0.92	0.57	0.62	6.91	76.97	4.34
CL/NW	0.98	0.83	0.85	4.24	26.37	4.46
CL/TA	0.41	0.13	0.32	0.38	0.53	1.13
CASH/S	0.05	0.13	2.75	6.62	57.31	8.34
CASH/TA	0.04	0.08	1.86	3.77	17.56	6.86
CA/S	0.59	0.22	0.38	3.70	23.56	3.50
QA/S	0.30	0.18	0.58	4.75	32.19	5.14

Table 3.10 Descriptive of distribution of financial ratios for survivors.

Financial ratio	Mean	Standard deivation	Coefficient of variation	Skewness	Kurtosis	K-S statistic
NI/S	0.03	0.15	4.78	-42.84	2135.03	15.86
FF/NW	0.23	0.17	0.73	0.47	8.15	4.34
FF/TA	0.11	0.09	0.80	-4.86	70.80	4.94
NI/TA	0.04	0.07	1.61	-7.46	125.79	8.26
NI/NW	0.09	0.14	1.53	-1.35	20.84	8.18
EBIT/S	0.06	0.16	2.78	-37.09	1759.38	12.91
EBIT/TA	0.08	0.09	1.22	-4.22	61.04	5.08
QA/TA	0.38	0.12	0.32	1.13	3.26	3.43
FF/S	0.09	0.16	1.89	-37.62	1782.28	13.55
CA/TA	0.69	0.12	0.18	-0.46	0.83	2.22
NW/S	0.43	0.29	0.67	-0.58	59.32	6.39
S/TA	1.41	0.74	0.52	10.96	196.23	9.84
TA/NW	2.15	1.01	0.47	2.70	17.70	9.71
FF/CL	0.33	0.30	0.92	-5.72	151.95	4.80
RE/TA	0.28	0.21	0.75	-5.94	87.34	4.92
CA/CL	1.92	0.75	0.39	2.31	10.02	6.59
QA/CL	1.04	0.53	0.51	3.44	21.69	8.55
CL/NW	0.95	0.83	0.87	3.31	19.55	9.74
CL/TA	0.40	0.16	0.39	4.52	78.88	3.44
CASH/S	0.05	0.12	2.36	12.35	270.46	18.31
CASH/TA	0.06	0.09	1.51	2.95	12.51	13.95
CA/S	0.55	0.25	0.46	11.38	254.23	8.45
QA/S	0.30	0.20	0.69	16.84	496.93	12.22

3.5 Principal Component Analysis

Many of the ratios included in the studies are highly correlated with one other. This overlapping occurs because the ratios are derived source accounting data. Such overlapping can still be found in most recent studies (Chen and Shimerda, 1981). For example the 56 items used in the computation of the 28 ratios included in the Elam (1975) study are derived from only 18 different pieces of financial data, and the 28 items for Deakin's (1972) ratios consist of only 10 separate pieces of data. The elimination of such overlapping would aid in the development of a useful set of financial ratios. Not all overlapping ratios, however, can be eliminated by visual inspection. Analysis of empirical relationships among financial ratios could be performed through correlation analysis (Gombola and Ketz, 1983). If two ratios are highly correlated, then the user could consider one of the pair to be redundant, discarding it with little loss of information. If two ratios are not highly correlated, then the user could consider each to measure a different aspect of company performance. Highly correlated ratios could be brought together into groups, where the groups would measure some different aspect of company performance. In this way the user could understand the relationships and patterns among the financial ratios in a variable set. Instead of grouping on the basis of the correlation coefficient, the grouping procedure could be performed via a statistical method designed to summarize such interrelationship, i.e., by using principal component analysis. Principal component analysis was developed by Harold Hotelling in the 1930's and has found extensive application in psychometrics and econometrics. One of the functions performed by principal component analysis is to group variables into a few components that retain a maximum of information contained in the original variable set. Its general objectives are (1) data reduction, and (2) interpretation. In this section principal component analysis is employed to isolate independent patterns of financial ratios. Principal component analysis, which employs financial ratios as variables and industrial companies as the cases, produces components of the financial ratios in terms of the industrial companies. The similarity of each variable in the reduced space with the components is measured by its component loadings. The reasons for choosing principal component analysis over other methods e.g. factor analysis which perform a similar function, is that it is a technique which may be applied to various types of data such as quantitative data and qualitative attributes either scored or scaled (Jeffers, 1978). In principal component analysis no assumptions are made about the form of the covariance or correlational structure of the variables. Factor analysis supposes that the data comes from a well-defined model where a set of underlying factors exist which account for the interrelationship of the variables, but not for their full variance. If the assumptions are not met, then factor analysis may give spurious results (Mardia et al 1979). In principal component analysis, the emphasis is on transformation from the observed variables to the principal components, whereas in factor analysis the emphasis is on the transformation from the underlying factor to the observed variables.

3.5.1 Procedure for A Principal Component Analysis

Our objective in this section is to construct uncorrelated linear combinations of the measured characteristics that account for much of the variation in the sample. The uncorrelated combinations with the largest variances will be called the principal components. Principal components are particular linear combinations of the p random variables $X_1, X_2, \ldots, X_p$. Geometrically these linear combinations represent the selection of a new coordinate system obtained by rotating the original system with $X_1, X_2, \ldots, X_p$ as the coordinate axes. The new axes represent the directions with maximum variability and provide a simpler and more parsimonious description of the covariance structure (Johnson and Wichern, 1982). Principal components depend on the covariance matrix S (or the correlation matrix r) of $X_1, X_2, \ldots, X_p$.

Let the random vector $X' = [X_1, X_2, \ldots, X_p]$ have the covariance matrix S with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$.

Consider the linear combinations

$$
\begin{aligned}
Z_1 &= a'_1 X = a_{11}X_1 + a_{21}X_2 + \ldots + a_{p1}X_p \\
Z_2 &= a'_2 X = a_{12}X_1 + a_{22}X_2 + \ldots + a_{p2}X_p \\
&\vdots \qquad\qquad\qquad \vdots \\
Z_p &= a'_p X = a_{1p}X_1 + a_{2p}X_2 + \ldots + a_{pp}X_p
\end{aligned}
\qquad \ldots(3.1)
$$

Then,

$$\mathrm{Var}(Z\ ) = a'\ Sa \qquad i = 1, 2, \ldots, p \qquad \ldots(3.2)$$

$$\mathrm{Cov}(Z_i, Z_j) = a'_i S a_j \qquad i, j = 1, 2, \ldots, p \qquad \ldots(3.3)$$

The principal components are those uncorrelated linear combinations $Z_1, Z_2, \ldots, Z_p$ whose variances in (3.2) are as large as possible.

The first principal component is the linear combination with maximum variance. That is, it maximizes $\mathrm{Var}(Z_1) = a'_1 S a_1$. It is clear that $\mathrm{Var}(Z_1) = a'_1 S a_1$ can be increased by multiplying any a_1 by some constant. To eliminate this indeterminacy it is convenient to restrict attention to coefficient vector of unit length. We therefore define

First principal component = linear combination $a'_1 X$ that maximizes $\mathrm{Var}(a'_1 X)$ subject to $a'_1 a_1 = 1$

Second principal component = linear combination $a'_2 X$ that maximizes $\mathrm{Var}(a'_2 X)$ subject to $a'_2 a_2 = 1$ and $\mathrm{Cov}(a'_1 X, a'_2 X) = 0$

At the ith step

ith principal component = linear combination $a'_i X$ that maximizes $\mathrm{Var}(a'_i X)$ subject to $a'_i a_i = 1$ and $\mathrm{Cov}(a'_i X, a'_j X) = 0$ for $j<i$

Principal component Z_p's are uncorrelated random variables which are linear functions of a correlated set of random variables $X_1, X_2, \ldots, X_p$, with the coefficients a_{ij}'s being the elements of the normalised eigenvectors of the correlation (or covariance) matrix of X_i's. The variances of the principal components are the eigenvalues of the matrix S. Assuming that the eigenvalues are ordered as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$, then λ_i corresponds to the ith principal component

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \ldots\ldots + a_{ip}X_p$$

$$\mathrm{Var}(Z_i) = \lambda_i .$$

An important property of the eigenvalues is that they add up to the sum of the diagonal elements (the trace) of S. that is

$$\lambda_1 + \lambda_2 + \ldots\ldots\ldots + \lambda_p = S_{11} + S_{22} + \ldots\ldots\ldots + S_{pp} .$$

Since S_{ii} is the variance of X_i and λ_i is the variance of Z_i, this means that the sum of the variances of principal components is equal to the sum of the variances of the original variables. Therefore in a sense the principal components account for all of the variation in the original data. The eigenvalues of the matrix S give the proportion of the total variability in the data explained by the individual components (the components are obtained from X_i's by an orthogonal transformation) and the largest eigenvalue gives the variance of the component which explains the largest variability in the data, the second largest eigenvalue gives the variance of the component which accounts for the maximum possible remaining variability etc. The coefficient a_{ij}'s are often refered to as component loadings. They indicate weighting of each variable and may be interpreted as the correlations between the principal components and the variables X_i's, provided the eigenvectors of the correlation matrix are scaled by the square root of corresponding eigenvalues (Morrison, 1976).

Principal component analysis is variable-sensitive: different components may be obtained if different sets of variables are fed into the principal component analysis.

The data were organised into a matrix consisting of rows (companies) and columns (financial ratios). The matrix was subjected to principal component analysis and the components were rotated using the varimax method in order to obtained subsets of specially related variables. Only the most important components having an eigenvalue greater than the average value of one (Kaiser's criteria) were retained. The components were interpreted in practice via the size and sign of the coefficients (loadings) of a component. The sizes indicate the correlations of variables with the respective component i.e. large component loading indicates that the component has highly significant correlation with respective variable, and furnish the basis for describing and naming these components. In this section the principal component analysis is used to reduce the dimensions of a data set from the number of variables (23) to a much smaller set of components.

The computer package SPSSX was used to carry out the analysis. Varimax rotated principal component analysis of the 23 ratio set used in the main analysis of the study was undertaken, and the rotated component loadings matrix for the analysis is shown in Table 3.11. Only the highly significant loadings (loadings > 0.50) are given in the Table, which also shows the eigenvalue, the percentage and cumulative percentage of variation accounted for by each component. The communalities also show that the proportion of variance of each variable accounted for by the six components is quite high, ranging from 64% to 95%. The number of components satisfying the criteria were six out of 23 variables, and when combined explained 83.3% of the original variance in the data.

Interpretion of the components may be described as follows:

Component one.

This component is an index which has high positive loadings on NI/S, FF/TA, NI/TA, EBIT/S, EBIT/TA, FF/S and FF/C.LIB. It explains 28.8% of the original variability in the data.

Component two.

Component two explains an additional 20.6% of the original variance. It is an index which has high positive loadings on FF/NW, FF/TA, NI/TA, NI/NW and EBIT/TA.

Component three.

This component is an index which has high positive loadings on QA/TA, QA/C.LIB, CASH/S, CASH/TA and QA/S. It explains a further 14.3% of the original variance.

Component four.

This component is a contrast of RE/TA, CA/C.LIB and QA/C.LIB against C.LIB/TA. It explains an additional 8.5% of the original variability.

Component five.

It explains an additional 6.8% of the original variance. This component is a contrast S/TA against NW/S and CA/S.

Component six.

This component is an index which has highly positive loadings on CA/TA, TA/NW and C.LIB/NW. It explains further 4.4% of the original variance.

The analysis of the groups (non-survivors and survivors) separately provided similar dimensions.

Table 3.11 Financial ratios and component loadings defining six financial ratio patterns for 445 non-surviving and surviving companies.

Variable	Components						Communa-lity
	1	2	3	4	5	6	
NI/S	0.97						0.95
FF/NW		0.90					0.86
FF/TA	0.51	0.75					0.90
NI/TA	0.57	0.67					0.89
NI/NW		0.86					0.84
EBIT/S	0.96						0.95
EBIT/TA	0.54	0.72					0.92
QA/TA			0.74				0.80
FF/S	0.94						0.93
CA/TA						0.51	0.66
NW/S					-0.64		0.82
S/TA					0.78		0.64
TA/NW						0.87	0.87
FF/C.LIB	0.65						0.81
RE/TA				0.64			0.67
CA/C.LIB				0.86			0.83
QA/C.LIB			0.58	0.68			0.84
C.LIB/NW						0.87	0.92
C.LIB/TA				-0.71			0.83
CASH/S			0.80				0.84
CASH/TA			0.88				0.80
CA/S					-0.66		0.82
QA/S			0.65				0.90
Eigenvalue	6.63	4.74	3.28	1.96	1.56	1.02	---
% of variance	28.8	20.6	14.3	8.5	6.8	4.4	---
Cum. % of variance	28.8	49.4	63.6	72.1	78.9	83.3	---

Conclusion

In this chapter, the structure of the data set has been discussed, and the small size of the sub-sample of bankrupt companies compared to other types of non-survival has been shown. This point is taken up again in Chapter 4, where a reclassification method is introduced. In addition, some of the properties of the financial ratio data have been described and, for both distributional properties and principal component, it has been found that assumptions of a single parent population for the survivors and non-survivors are appropriate.

CHAPTER FOUR

DISCRIMINANT ANALYSIS AND CLASSIFICATION OF THE "MERGED" AND "OTHER" COMPANIES

4.1 Introduction

Discriminant analysis is a multivariate technique concerned with separating distinct sets of objects (or observations) and with allocating new objects (or observations) to previously defined groups. The discriminant function may be accepted as the explicitly devised method of classification research. The procedure was developed in 1936 by Fisher to answer perhaps one of the most fundamental of all systematic problems of the taxonomic variety; it stands as both the first clear statement of the problem of discrimination and the first proposed solution (Al-Moswie, 1982).

In this chapter we shall introduce linear and quadratic discriminant functions and their applications to forecasting company failure. We shall also introduce methods to reclassify the "merged" and "other" companies into either surviving or non-surviving companies using two procedures: stepwise discriminant analysis and survival models.

4.2 Linear Discriminant Analysis

Various authors have used discriminant analysis to classify companies as either surviving or non-surviving on the basis of financial ratios (e.g. Altman, 1968 and Barnes, 1990). Detailed reviews of the applications were discussed in Chapter 2.

The first task of a linear discriminant analysis is to select a set of variables X_1, X_2,.....,X_K that best discriminate between, or separate, groups e.g. non-surviving or surviving companies. The variables measure attributes on which the groups differ to some extent, otherwise the groups cannot be distinguished by means of the X's alone. Also the groups should be partly overlaping, otherwise discrimination is not necessary. The object of linear discriminant analysis is then to find a linear function in the X's so that as many cases as possible belonging to the first group lie on one side of the function and as many cases as possible belonging to the second group lie on the other side. The allocation criteria is based on the likelihood of a case belonging to a group with a boundary where the likelihoods are equal (Kendall, 1980). We therefore seek a new variable Z such that

$$Z = a_1X_1 + a_2X_2 + \ldots\ldots\ldots + a_kX_K \qquad \ldots\ldots\ldots(4.1)$$

where a_1, $a_2, \ldots\ldots, a_k$ are the coefficients of the discriminant function Z.

In the case of 2 groups the mean values are

$$\bar{Z}_1 = a_1\bar{X}_{11} + a_2\bar{X}_{12} + \ldots\ldots\ldots\ldots + a_k\bar{X}_{1K}$$

where $\bar{X}_{ir}$ is the mean of the r^{th} measurement of the i^{th} group, $i = 1, 2$; $r = 1, 2, \ldots\ldots, k$.

And the difference between the means of the two groups is

$$D = (\bar{Z}_1 - \bar{Z}_2) = a_1d_1 + a_2d_2 + \ldots\ldots\ldots + a_kd_k \qquad \ldots.(4.2)$$

where

$$d_r = \bar{X}_{1r} - \bar{X}_{2r} \qquad , \qquad r = 1,\ 2,\ \ldots\ldots, k.$$

Now the coefficient of the optimum linear discriminant function should be chosen to maximize the following function (Johnson and Wichern, 1982):

$$G = \frac{(\bar{Z}_1 - \bar{Z}_2)^2}{\sum_{i=1}^{2} \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2} \qquad \ldots\ldots\ldots\ldots\ldots (4.3)$$

where Z_{ij} denotes the Z value of the j^{th} individual in the i^{th} group where i = 1 or 2.

Let S_{pq} be

$$S_{pq} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (X_{pij} - \bar{X}_{pi})(X_{qij} - \bar{X}_{qi})$$

where

$$p,\ q = 1,\ 2,\ \ldots\ldots\ldots\ldots,\ k$$

Then

$$\sum_{i=1}^{2} \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2 = \sum_{p,q=1}^{k} a_p a_q S_{pq}$$

and

$$(\bar{Z}_1 - \bar{Z}_2)^2 = \sum_{p,q=1}^{k} a_p a_q d_p d_q$$

Now G defined in (4.3) may be written as :

$$G = \frac{\sum_{p,q=1}^{k} a_p a_q d_p d_q}{\sum_{p,q=1}^{k} a_p a_q S_{pq}} \qquad \ldots\ldots\ldots\ldots(4.4)$$

Since the a's are to be determined such that G is a maximum, it is necessary that $\partial G/\partial a_r = 0$,

for r = 1, 2, ..., k at the maximizing point.

Then

$$a_1S_{r1} + a_2S_{r2} + \ldots\ldots + a_kS_{rk} = \frac{a_1d_1 + a_2d_2 + \ldots + a_kd_k}{G}$$

$$\ldots\ldots\ldots(4.5)$$

Now $(a_1d_1 + a_2d_2 + \ldots. + a_kd_k)/ G$ is independent of r, and it could be considered as a constant C.

Hence (4.5) may be written as :

$$a_1S_{r1} + a_2S_{r2} + \ldots\ldots + a_kS_{rk} = Cd_r \qquad \ldots\ldots\ldots(4.6)$$

Let C = 1, then (4.6) may be written as :

$$a_1S_{r1} + a_2S_{r2} + \ldots\ldots + a_kS_{rk} = d_r \qquad \ldots\ldots\ldots(4.7)$$

This formula gives the values of a's. Then we substitute in (4.1) to find the discriminant function Z.

Let us consider the special case of $K = 2$ in (4.1). This means, for example, that the function consists of two financial ratios as classification variables. Here X_1 may be a measure for liquidity and X_2 a measure for the profitability of the company. Now suppose that the coefficients a_1 and a_2 are known and that their signs are positive. If we now have the disposal of the values of X_1 and X_2 of a company then we can calculate from the function the Z-score of that company. If that score is high then the company is classified as a surviving one and as a non-surviving if that score is low, since the liquidity and the profitability of surviving companies on average will be high and those of the non-surviving will be low. The classification procedure makes a comparison of the Z-score of a company with a critical score, say Z^*, such that :

$$\begin{aligned} \text{if } Z > Z^*, &\quad \text{then company} \longrightarrow \text{in the surviving group} \\ \text{while if } Z < Z^*, &\quad \text{then company} \longrightarrow \text{in the non-surviving group.} \end{aligned}$$

Of course the classification will not proceed without error. There are two possible types of error:

(a) incorrectly classifying a company into the non-surviving group (type II error)

(b) incorrectly classifying a company into the surviving group (type I error).

Although it is desirable to minimise both type I and type II errors, the former may be considered to be the more important one. If an investor were to buy stock with the guidance of a discrimination model, he would presumably prefer his investment to be on the safe side. The misclassification of a surviving company as non-surviving has less serious consequences than the classification of a non-surviving as healthy. This point should be considered in the development of a satisfactory model for predictive classification.

Note that the justification for the technique is based on likelihood that the data come from multivariate normal distributions with the same variance-covariance matrix in both populations. If this is not the case, then it may still be an intuitively reasonable technique to use (Manly 1986 , Srivastava and Carter 1983, Betts and Belhoul 1982). Fisher's procedure, for example, did not depend on the form of the parent populations, apart from the requirement of identical covariance structures. Studies by Krzarowski (1977) and Lachenbruch (1975) have shown, however, that there are non-normal cases where Fisher's linear classification function performs poorly even though the population variance-covariance matrices are the same.

As indicated above, the linear discriminant function depends, for ensuring minimization of the probability of misclassification, on the assumptions of separate multivariate normality for the populations in the analysis and equality of dispersion matrices. If however normality holds, but as in the case of Lachenbruch (1975), the variance-covariance matrices are heterogeneous, then the theory suggests fitting a quadratic discriminant functions. It should be remembered that equality of dispersion matrices is conventionally tested prior to the fitting of a quadratic function, and that the appropriate Bartlett-Box criterion is sensitive to departures from the assumption of multivaiate normality. Therefore the effects of departure from normality in the two-group discriminant case and fitting a quadratic model in the case of unequal variance-covariance matrices may, depending on the type of non-normality, well make matters worse than the use of a linear approach (Taffler, 1982). This point is illustrated in the next sections. Some fairly well known problems associated with multivariate discriminant analysis may be avoided by using conditional logit analysis which is discussed in detail in Chapter 5.

4.3 Stepwise Discriminant Analysis

Stepwise discriminant analysis is used to select a relatively small subset of variables that would contain almost as much information as the original collection. In this procedure some variables are selected as being best for classification and the remaining ones are discarded. This procedure is usually available as a computer program in the form of a forward selection technique that adds variables one by one depending on the discriminating ability of each and may be regarded as the most efficient procedure in that the most important variables are selected first. Stepwise discriminant analysis has the advantage of being readily available in all major statistical packages and the procedure can sift through a large number of variables and indicates those most promising for classification.

The particular program used in this study is the one in the SPSSX statistical package. This program at each step in the forward process calculates the Wilk's lambda statistic and F-ratio for all the variables under consideration. This statistic measures the discriminating power gained (or lost) by adding (or dropping) a variable. Then the variable that has the smallest lambda statistic, i.e. adds the most to the discriminating ability, is added.

4.4 Quadratic Discriminant Function

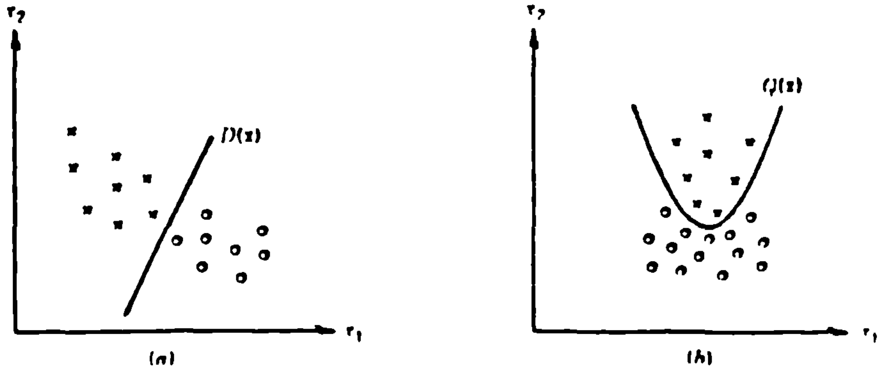
The assumption of equal variance-covariance matrices in linear discriminant analysis is rarely satisfied in practice, although in some cases the matrices are close enough that it makes little or no difference to the results to assume their equality (Lachenbruch, 1975). If the dispersion matrices are different then a quadratic discriminant function would be more appropriate. For example in Figure 4.1a, $S_1 = S_2$, where

$$S_1 = \frac{1}{n_1-1} \sum (X_{1j} - \bar{X}_1)^2$$

$$S_2 = \frac{1}{n_2-1} \sum (X_{2j} - \bar{X}_2)^2$$

so that the two sets of points, which are similar in shape, can be adequately separated by a straight line. However, in Figure 4.1b, S_1 does not equal S_2 so that one set of points is long and thin vertically while the other is circular. Here a curved discriminant function gives adequate separation.

Figure 4.1 Use of (a) linear discrimination D(x) and (b) quadratic discrimination Q(x) to separate two sets of points.



Consider the following multivariate normal densities with unequal covariance matrices ($S_1 \neq S_2$):

$$f_i(X) = \frac{1}{(2\pi)^{p/2}|S_i|^{1/2}} \exp\left[-\frac{1}{2}(X - \bar{X}_i)' S_i^{-1}(X - \bar{X}_i)\right]$$

for i = 1, 2.

Then the quadratic discriminant function Q(X) is

$$Q(X) = \log\left[\frac{f_1(X)}{f_2(X)}\right] \qquad \ldots(4.8)$$

$$= \frac{1}{2}\log\left(\frac{|S_1|}{|S_2|}\right) - \frac{1}{2}(X - \bar{X}_1)' S_1^{-1}(X - \bar{X}_1)$$

$$+ \frac{1}{2}(X - \bar{X}_2)' S_2^{-1}(X - \bar{X}_2)$$

In this case, the quadratic classification rule applied is as follows (Johnson and Wichern 1982, Seber 1984): allocate to population 1 if

$$-\frac{1}{2} X'(S_1^{-1} - S_2^{-1})X + (\bar{X}_1' S_1^{-1} - \bar{X}_2' S_2^{-1})X \geq K$$

where

$$K = \frac{1}{2}\log\left(\frac{|S_1|}{|S_2|}\right) + \frac{1}{2}(\bar{X}_1' S^{-1}\bar{X}_1 - \bar{X}_2' S_2^{-1}\bar{X}_2)$$

and allocate to population 2 otherwise.

4.5 Application of Discriminant Analysis

This section contains the results of applying the discriminant analysis technique to the financial data base that was used in this study. There are problems associated with discriminant analysis which are dealt with here. If the data has a sample bias, as in this case where there are approximately 3000 observations for the surviving group of companies and only approximately 600 observations for the non-surviving group for the entire period under consideration, an unacceptable conclusion may be drawn from the analysis. The discriminant model may have an apparently good classification percentage but this may stem from the sample bias (Zmijewski, 1984). For example if most of the surviving companies are correctly classified but most of the non-survivors are not, the classification percentage will still be high. A method of resolving this may be to pair the non-survivors with survivors by using equal prior probabilities and then perform the analysis. This process will be investigated.

Discriminant analysis is used to carry out the analysis for 23 financial ratios *before* reclassification of the "merged" and "other" companies. In this respect, two cases of discriminant analysis are investigated:

a(1)- The bankrupt companies alone were taken as a first group (non-surviving group) and surviving companies as a second group. The reporting period was from 1971-1984 and all the data for this period was included. The sample sizes were n_1=162 and n_2=2997 respectively.

a(2)- The bankrupt, "merged" and "other" companies were formed into a single group (non-surviving group) and surviving companies as a second group. The reporting period was taken to be the same, from 1971-1984 and again the data for the entire period was included. The sample sizes were n_1=640 and n_2=2997 respectively.

For the above two cases (i.e. a(1 and 2)), the discriminant method is investigated for each of the five years prior to failure.

The computer packages SPSSX and SAS were used to carry out the following discriminant runs *before* reclassification of "merged" and "other" companies:

(i) Two groups (non-surviving and surviving), using a stepwise method and setting prior probability to 'sample size' during classification, i.e. using a population based sample where the number of survivors far exceeded the number of non-survivors.

(ii) Same as (i) except for using a paired sample i.e. setting prior probability of 1:1.

Here the groups were assigned the value 1 for non-surviving companies and the value 2 for the surviving companies.

The results of the analyses of each of the two runs are given in Tables 4.1 and 4.2, *before* reclassification of "merged" and "other" companies. It can be seen from Table 4.1(i) that equal prior probabilities gives the best discriminant function in the sense that its accuracy is the best among the two runs. Note that for (ii) the type I error is as high as 88.1% while type II error is as low as 1.2% but the classification accuracy (95.3%) appear to be good overall, so with respect to type I error the model is extremely poor. A possible reason for this is that there were many surviving companies, and the non-surviving company data appeared as noise when compared to the survivors . These results were to be expected and concur with the comments made by Zmijewski (1984). For (i), type I error is 15.5% and type II error 18%. We may therefore assume that the function obtained from (i) is the best function and the percentage of cases correctly classified was 82.3%. This indicates that the model developed misclassifies only 17.7% of the total number of observations. But, in the case of a quadratic discriminant function, we found that the percent of cases correctly classified was lower at 74.1% and at the same time the 2 types of error were higher at 38.1% and 25.1% respectively. Stepwise discriminat analysis shows that 12 variables out of 23 contributed significantly to the discriminant function. These are, in descending order of importance: Net worth/sales (NW/S), Funds flow/current liabilities (FF/C.LIB), Current liabilities/total assets (C.LIB/TA), Current assets/ total assets (CA/TA), Current assets/current liabilities (CA/C.LIB), Net income/net worth (NI/NW), Quick assets/current liabilities (QA/C.LIB), Retained earnings/total assets (RE/TA), EBIT(earning before interest & tax)/sales (EBIT/S), Funds flow/sales (FF/S), Funds flow/total assets (FF/TA), and Current assets/ sales (CA/S).

The model developed in Table 4.2 with the bankrupt, "merged" and "other" companies, treated as non-surviving group of companies and surviving companies as a second group (i.e. case a(2)). It can be seen that case (i) with equal prior probabilities gives the best linear discriminant function in the sense that its accuracy is the best among the two cases, with type I error being 19.5% and type II error 22.3% and the percent of cases correctly classified was 78.2%. This indicates that the model developed shows that the misclassified comprise only 21.9% of the total number of observations. Again, for case (ii), where nearly all the observations on surviving companies are correctly classified, whilst there are few correct classifications for the non-survivors, confirms the problem of sample bias (Zmijewski, 1984). But, in the case of a quadratic discriminant function, we found that the percent of cases correctly classified was lower at 64.9% and type I and type II errors were higher of 33.7% and 34.1% respectively .

Stepwise discriminant analysis showed only 13 variables out of 23 to be significant. These were in descending order of discriminatry power: Quick assets/total assets (QA/TA), Current liabilities/ total assets (C.LIB/TA), Funds flow/current liabilities (FF/C.LIB), EBIT/sales (EBIT/S), Retained earnings/total assets (RE/TA), Current assets/ current liabilities (CA/C.LIB), Sales/total assets (S/TA), Net worth/sales (NW/S), Funds flow/total assets (FF/TA), Funds flow/sales (FF/S), Total assets/net worth (TA/NW), Quick assets/current liabilities (QA/C.LIB) and Net income/net worth (NI/NW).

Table 4.1 Results of linear discriminant analysis *before* reclassification of "merged" and "other" companies for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

Actual group		Method	prior prob.	Predicted group membership		percentage correctly classified
				1	2	
(i)	non-survivor (1)	stepwise (Wilks)	equal (1:1)	84.5%	15.5%	82.3%
	survivor (2)			18%	82%	
(ii)	non-survivor (1)	stepwise (Wilks)	sample size	11.9%	88.1%	95.3%
	survivor (2)			1.2%	98.8%	

Table 4.2 Results of linear discriminant analysis *before* reclassification of "merged" and "others" companies for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

Actual group		Method	prior prob.	Predicted group membership		percentage correctly classified
				1	2	
(i)	non-survivor (1)	stepwise (Wilks)	equal (1:1)	80.5%	19.5%	78.2%
	survivor (2)			22.3%	77.7%	
(ii)	non-survivor (1)	stepwise (Wilks)	sample size	3.3%	96.7%	84.9%
	survivor (2)			0.4%	99.6%	

Analysis for 5 years prior to failure

As indicated above, the main issue raised by moving from a matched sampling basis is that the structure of the survivor group no longer reflects that of the non-survivors. The survivor group now contains all companies which have not yet failed (i.e. in this data set, 359 survivors and 95 non-survivors, of which 21 were bankrupts). Also, given the small number of listed companies failing each year, the need to generate sufficiently large samples of non-survivors by including companies which failed at different points in time produces further problems of sample structure.

For, the first analysis (see Table 4.3a and 4.4a), the survivor group was censored at the date of the last entry in the data set. In this case, given that most survivors reported financial results in 1984, the 'one year before' vector was consequently the 1983 data, and so on.

However, a preferable approach would be to apply stratified sampling in order to select randomly the survivors such that the failures in any given year when expressed as a proportion of the total number of non-survivors is reflected in the censored group. For instance, the number of companies failing in 1978, for example, (5 companies) as a proportion of the total (21) was applied to the number of survivors (337) resulting in a stratified sample (80 companies selected randomly from those survivors which were in existence in 1978 and had reported in 1977) whereby the 'one year before' data was the 1977 results, the 'two years before' data was the 1976 results, and so on. Then, the procedure was repeated for companies failing in 1979, where the 'n years before' data was the 1979-n reported results. This was repeated for all years in which there had been failures. With this procedure, the sample for the final year comprises those survivors which had not been selected randomly for the other strata. This approach is referred to here (and, as far as can be seen, the technique is not known elsewhere) as "randomly-censored stratified sampling". The stratified samples were as follows:

Years	Number of bankruptcies	Randomly-censored stratified sample
1983	7	113
1982	4	64
1981	2	32
1980	1	16
1979	2	32
1978	5	80
Total	21	337

Although the number of companies with data n years before random censoring varies (as newly listed companies with 2 years or more data were included), the structure for the 5 years may be illustrated as follows:

Example of sample structure using "randomly-censored stratified sampling"

	Number of companies Years before failure or censoring				
	1	2	3	4	5
bankrupt	21	18	17	17	14
survivor*	337	305	262	229	206

* average of 32 runs

It should be noted that the data set was reduced to 337 companies by removing (a) all companies reporting for 1984 only (i.e. 9 new listing in 1984) and (b) those which were survivors but had not reported in 1984 (13 companies). It can be seen that the survivor group now reflects the proportions and the structure of the bankrupt group.
The procedure was repeated 32 times, and applied to the case of bankrupt companies (a(1)), of all non-survivors (a(2)) (see Table 4.3b and 4.4b), and of the reclassified grouping where companies taken over but which had a high likelihood of failure were reassigned to the bankrupt group (see Section 4.6.2 and Table 4.8b).

The results of applying a linear discriminant analysis by using randomly-censored stratified sampling technique reduced type I error in all cases in Table 4.3b compared with Table 4.3a, but in Table 4.4b increased the type I error in some cases (e.g. 1 and 3 years prior to failure) compared with Table 4.4a.

It can be seen from Tables 4.3(a) and (b) and 4.4(a) and (b) that applying a linear discriminant analysis to years prior to failure increased the type I error substantially and type II error, compared to the results in Tables 4.1 and 4.2 for the entire set of data.

Compared to published studies where in most cases the sample sizes for the two groups were similar, the linear discriminant model classification results are not good. The high misclassification errors may be due largely to the fact that the sample sizes for the two groups of companies are quite different. Nevertheless, note that Shailer's (1990) model, using 33 non-surviving to 39 surviving companies gives type I error (39.4%) and type II error (28.2%), and, in spite of using equal sample size for both surviving and non-surviving companies, the results were poor. Taffler's (1982) model, using 23 non-surviving to 45 surviving companies gives only the type I error (4.3%) and no type II error on one year prior to failure data. Also, Tisshaw's (1976) model using 31 non-surviving to 62 surviving companies gives type I error (3.2%) and type II error (1.6%). However, both of these used "solvent" companies as survivors, and it is therefore to be expected that their models performed better. However, a model by Luk (1984) using 27 non-surviving to 170 surviving companies gives type I error (16%) and type II error (19.6) on one year prior to failure data.

Nevertheless, it should be noted that the interest in developing a discriminant analysis model here is in order to compare the results obtained with those from logit analysis and survival analysis.

The issue to be emphasised here is that the poor results are a reflection of the application of the linear discriminant model using a realistic (i.e. unbalanced) sample. This is a major shortcoming of conventional modelling procedures, where sample bias is a feature of the data, and this issue is taken up later in the thesis when alternative modelling approaches are introduced.

Table 4.3 Results of linear discriminant analysis *before* reclassification of "merged" and "other" companies for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) for five years prior to failure by using equal prior probabilities.

(a) - survivor data: one year prior to censoring

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error (%)	33.9	61.1	52.9	64.7	50
Type II error (%)	17.1	20.2	26.7	6.5	35
Correctly classified (%)	81.9	77.7	72	90.6	64.4
No. of cases non-survivor	21	18	17	17	14
survivor	350	336	326	320	320

(b) -survivor data: randomly-censored stratified samples

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error(1) (%)	24.5 (13.4)	50.9 (11.4)	42.3 (11)	63.9 (3.6)	49.1 (6.7)
Type II error(1) (%)	25.4 (2.6)	19.8 (3.2)	18.4 (2.6)	9.3 (3.5)	25.9 (8.8)
Correctly(1) classified (%)	74.9 (2.3)	78.5 (2.5)	80.2 (2.1)	87 (2.9)	72.6 (8.1)
No. of cases non-survivor	21	18	17	17	14
survivor(1)	337	305	262	229	206

(1) All the values are averages based on 32 runs.

Figures in parenthesis are standard deviations.

Table 4.4 Results of linear discriminant analysis *before* reclassification of merged and "other" companies for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) for five years prior to failure by using equal prior probabilities.

(a) - survivor data: one year prior to censoring

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error (%)	35.8	38.5	31.9	51.6	49
Type II error (%)	30.9	31.5	27.6	22.7	29.1
Correctly classified (%)	68.1	67.2	71.7	72.5	68.2
No. of cases non-survivor	95	78	69	64	51
survivor	350	336	326	320	320

(b) -survivor data: randomly-censored stratified samples

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error(1) (%)	41.8 (6.3)	36.8 (4.8)	37.5 (4)	45.8 (5.2)	46 (5.4)
Type II error(1) (%)	34.7 (7.5)	39.3 (3.5)	40.9 (2.8)	28.3 (2.9)	38.5 (6.8)
Correctly(1) classified (%)	63.8 (5.6)	61.2 (2.5)	59.7 (2)	68.2 (2.1)	60.2 (4.7)
No. of cases non-survivor	95	78	69	64	51
survivor(1)	337	314	283	255	220

(1) All the values are averages based on 32 runs.

Figures in parenthesis are standard deviations.

4.6 Reclassification of the "Merged" and "Other" Companies

As noted earlier, previous prediction models have defined failure narrowly, mainly as bankruptcy. Altman (1968) and Ohlson (1980) restricted their definition of failure to companies which have filed bankruptcy petitions under Chapter X or XI of the U.S. Federal Bankruptcy Act. On the other hand, Beaver (1966) defined failure as the inability of a company to pay its financial obligations as they mature, but even in his case 75% of his sample companies were in the bankruptcy category. Likewise, in the case of Blum (1974), who defined failure as events signifying an inability to pay debts as they fall due, filing for bankruptcy or making an agreement with creditors to reduce debts, 90% of his cases were bankrupt.

This narrow definition of failure led to the restricted sample sizes used in the above studies. In turn their reliability and potential contribution are restricted. On the other hand, there would be three advantages in defining failure broadly to include an expanded set of events on the continuum (Lau, 1982). First, many companies recover after getting into the earlier stages of financial distress and avoid eventual failure, but it is still desirable to identify such companies in advance. Second, a broader definition of failure enables more companies to be included in the analysis sample used for constructing the prediction models, and the resultant models should benefit from this additional information on companies in the different stages of the failing continuum. Third, predicting a wider range of failing events would broaden the applicability of the prediction model for analysts and decision makers.

Furthermore, as mentioned earlier, companies are created and they may fail, either through liquidation or bankruptcy or through merger or acquisition or for some other reasons. Some survive, and an alternative perspective is that at any given point in time, the survivors are those companies which have not yet been taken over or have not yet failed. It can be argued that bankruptcy is not the only yardstick of failure. Various companies at various times go through a period of financial instability which, if remedial steps are not taken, may lead to bankruptcy. Companies which merge to avoid bankruptcy may also be considered as failures, as mentioned in Chapter 3 Section 3.1. Most of the previous studies do not mention the acquisition of companies and some of them do not define "failure" (Dambolena and Khoury 1980, Keasey and Watson 1987), but just use the term "failed companies". In our case we have 74 (merged, taken over and other) such companies in comparing with only 21 bankrupt companies. We intend in this section to utilize the information and increase the sample size particularly the bankrupt companies by using two procedures: stepwise discriminant model and survival model methods. Even though the survival model is given in detail in a later chapter, it is used here in order to describe a method for reclassifying "merged" and "other" companies.

standard comparison

4.6.1 Stepwise Discriminant Method of Reclassification

A stepwise discriminant model which we discussed in Sections 4.2 and 4.3 may be used to reclassify the "merged" and "other" companies with 23 variables (financial ratios) by using bankrupt companies as a single group, treated as non-surviving, and using surviving companies as a second group (i.e. case a(1) in Section 4.5). The results of this analysis were given in Table 4.1 of Section 4.5. By using the final linear discriminant model to classify the "merged" and "other" companies, of which there are 74, we found that only five companies (those numbered 27, 39, 54, 56 and 65) were classified as bankrupt and the rest, 69 companies, were classified as live. The analysis was then re-run with two new groups: bankrupt plus 5 others (treated as the non-surviving group) and surviving plus 69 others (treated as the surviving group). The results of the analysis are given in Table 4.5. These show that the percentage correctly classified was 80.4% and type I error was 45% and type II error 14.3%. But, in the case of a quadratic discriminant analysis, we found that the percent of cases correctly classified was 66.3% and type I error was 63.1% and type II error 27.5%. Stepwise discriminant analysis showed that only 8 variables out of the 23 were significant. These, were in descending order of importance, Net worth/sales (NW/S), Current liabilities/total assets (C.LIB/TA), Funds flow/sales (FF/S), Currents assets/current liabilities (CA/C.LIB), Current assets/total assets (CA/TA), Quick assets/ current liabilities (QA/C.LIB), Current assets/ sales (CA/S) and EBIT/total assets (EBIT/TA).

Basically, the stepwise discriminant method for reclassification in this case gives a lower percentage of correctly classified cases and the type I error is larger than when these companies were left out, as we have seen in Section 4.5. These results indicate that using stepwise discriminant method for reclassification is not powerful enough.

Table 4.5 Results of linear discriminnt analysis *after* reclassification of "merged" and "other" companies by using bankrupt and 5 others companies as a non-surviving group and the surviving and 69 others companies as surviving group.

Actual group	Method	prior prob.	Predicted group membership		percentage correctly classified
			1	2	
non-survivor (1)	stepwise (Wilks)	equal (1:1)	55%	45%	80.4%
survivor (2)			14.3%	85.7%	

4.6.2 An Alternative Method of Reclassification

The stepwise discriminant method is compared here with an alternative based on survival analysis. As discussed later in Chapter 6, survival analysis is concerned with a population of companies where for some companies we may also observe their time to "loss" from the study, or censoring. For a company which is censored, we know only that the time to failure is greater than the censoring time (see Chapter 6 for more details). For such a company, the time to failure is a random variable, lifetime is denoted by T and the probability of a company surviving to time t defines a survival function as

$$S(t) = pr\ (T \geq t)$$

which is a nonincreasing function of t. The underling idea in survival analysis is that of hazard function which gives the conditional failure rate. It is defined as the probability of failure during a small time interval $[t, t+\Delta t]$, assuming that the individual has survived to the beginning of the interval, or as the limit of the probability that an individual fails in a small interval, given that the individual has survived to time t. Survival model is used here to classify the "merged" and "others" companies according to the maximum likelihood (M.L.E) principle for Weibull regression model with 12 variables (selected from stepwise regression by using bankrupt companies as a non-surviving group and surviving companies as a surviving group, we obtain 12 variables which were the same 12 variables obtained from stepwise discriminant analysis above). The procedure is that for each individual in the sample we observe the vector of explanatory variables X_i and a pair of variables (t_i, δ_i). The censoring indicator δ_i takes the value 1 if the survival time t_i for the ith observation is uncensored, and zero if it is censored.

The likelihood function is (Altkin, et al 1989)

$$L\ (\beta,\gamma) = \prod_{i}^{n}\ [f(t_i)]^{\delta_i}\ [S(t_i)]^{1-\delta_i}$$

$$= \prod_{i}^{n}\ [h(t_i)]^{\delta_i}\ S(t_i) \qquad \ldots\ldots\ldots\ldots\ldots(4.9)$$

where

$h(t_i) = \dfrac{f(t_i)}{S(t_i)}$, $S(t_i)$ and $f(t_i)$ are the hazard,

survival and density functions respectively.

Assuming a constant shape parameter, we have for the the Weibull regression model,

$$h(t;X) = \lambda\gamma(\lambda t)^{\gamma-1}\ e^{X\beta} \quad , \quad (\ t\geq 0\ ;\ \lambda,\ \gamma>0) \qquad \ldots\ldots\ldots(4.10)$$

$$S(t;X) = \exp\ [-(\lambda t)^{\gamma}\ \ e^{X\beta}\ ] \qquad \ldots\ldots\ldots(4.11)$$

where

λ = scale parameter

γ = shape parameter

X = regression vector

β = coefficient of regression vector writing

$\theta = (\lambda t)^{\gamma} e^{X\beta} = H(t)$

(where H is the integrated hazard function.)

We obtain

$$L(\beta,\gamma) = \prod_{i=1}^{n} [\gamma\theta_i / t_i]^{\delta_i} e^{-\theta_i} \qquad \text{...........(4.12)}$$

The probability density function of the log failure time Y for Weibull regression model is (Kalbfleisch and Prentice 1980)

$$\sigma^{-1} \exp\left[\frac{Y - X\beta}{\sigma} - \exp\left(\frac{Y - X\beta}{\sigma}\right)\right] \quad , \ -\infty < Y < \infty \quad \text{....(4.13)}$$

where

$\sigma = \gamma^{-1}$

Model (4.13), may be written as

$$Y = X\beta + \sigma W \qquad \text{........(4.14)}$$

where W has a standard extreme value distribution with p.d.f.

$$\exp(W - e^{W}) \quad , \quad -\infty < W < \infty ,$$

and the likelihood function may then be written as

$$L(\beta,\sigma) = \prod_{i}^{n} [\sigma^{-1} f(W_i)]^{\delta_i} [S(W_i)]^{1-\delta_i} \qquad \text{.........(4.15)}$$

where

$W_i = (Y_i - X_i\beta) / \sigma$

The maximum likelihood equations $\frac{\partial \log L}{\partial \beta} = 0$ are readily solved by the Newton-Rapson method to obtain the M.L.E. $\hat{\beta}$

In order to estimate the likelihood that an unclassified company that has been acquired or wound-up, for reasons other than bankruptcy possesses characteristics similar to those of bankrupt companies or, alternatively, surviving companies, we used the Weibull survival likelihood function to estimate the log-likelihood for the known bankrupt companies and for the known survivors (i.e case a(1) in Section 4.5). We do this by adding each unclassified company first to the bankrupt group and then to the surviving group, and recompute the separate log-likelihoods and observe the difference. The results are shown in Table 4.6 and plotted in Figure 4.2. An increase in the log-likelihood indicates a better fit of the model to the data. Only in a few cases is there an observable change in the log-likelihood for the surviving companies. This is mainly due to the large sample size of this group. But, when added to the smaller sample of bankrupt companies 34 unclassified companies cause a significant decrease in the log-likelihood, leaving 40 companies which we may be deemed to have characteristics similar to the known bankrupt companies. These results are infered from the fact that a decrease in the deviance (= -2(log-likelihood), which has χ^2 distribution) of greater than 4 causes a significant reduction in the goodness-of-fit of the model at the 5% level. Figure 4.3 gives a plot of the difference between the deviances when all 74 unknown companies are reclassified.

Table 4.6 The results of changes in the log-likelihood when the 74 unclassified companies are added individually to the bankrupt and surviving groups for the purpose of reclassification.

Company	Log-likelihood when unclassified company added to bankrupt group	Log-likelihood when unclassified company added to surviving group	Difference in log-likelihood
1	-63.2652	-61.4044	1.8608
2	-63.3236	-61.3816	1.9420
3	-65.7885	-61.3715	4.4170
4	-66.5678	-61.3681	5.1997
5	-67.3547	-61.3673	5.9874
6	-63.3108	-61.3739	1.9369
7	-66.6307	-61.3687	5.2620
8	-63.1955	-61.3767	1.8188
9	-63.2899	-61.3889	1.9010
10	-62.0117	-61.7141	0.2976
11	-62.8826	-61.4869	1.3957
12	-63.0057	-61.4678	1.5379
13	-63.0199	-61.4649	1.5550
14	-67.5485	-61.1813	6.3672
15	-63.1802	-61.3939	1.7863
16	-65.5911	-61.3735	4.2176
17	-65.6185	-61.3724	4.2461
18	-63.2023	-61.3799	1.8224
19	-63.3100	-61.3865	1.9235
20	-69.6008	-61.3667	8.2341
21	-67.1846	-61.3677	5.8169
22	-70.4302	-61.3666	9.0636
23	-63.3282	-61.3784	1.9498
24	-63.1513	-61.4396	1.7117
25	-61.8594	-61.8461	0.0133
26	-63.0282	-61.4245	1.6037
27	-61.7555	-61.1752	-0.4197
28	-68.1707	-61.3670	6.8037
29	-67.6412	-61.3668	6.2744
30	-68.4578	-61.3668	7.0910
31	-63.0991	-61.3777	1.7214
32	-66.6806	-61.3688	5.3118
33	-63.1383	-61.3831	1.7552
34	-62.9450	-61.4230	1.5220
35	-67.1975	-61.3671	5.8304
36	-65.7149	-61.3692	4.3457
37	-66.4691	-61.3691	5.1000
38	-63.1573	-61.3772	1.7801
39	-62.3260	-62.7117	-0.3857
40	-67.4822	-61.3674	6.1148
41	-63.3008	-61.3790	1.9218
42	-67.7658	-61.3672	6.3986
43	-65.7591	-61.3711	4.3880

44	-63.2930	-61.4386	1.8544
45	-63.0095	-61.4633	1.5462
46	-67.3932	-61.3676	6.0256
47	-66.0469	-61.3707	4.6762
48	-63.3068	-61.4108	1.8960
49	-66.2224	-61.3691	4.8533
50	-63.0093	-61.3778	1.6315
51	-68.4832	-61.3668	7.1164
52	-71.6304	-61.3666	10.2638
53	-65.4528	-61.3741	4.0787
54	-62.1213	-61.4033	0.7180
55	-63.0961	-61.3836	1.7125
56	-61.8944	-62.8501	-0.9557
57	-63.1708	-61.3774	1.7934
58	-68.2079	-61.3670	6.8409
59	-69.6237	-61.3667	8.2570
60	-66.4381	-61.3687	5.0694
61	-66.3068	-61.3692	4.9376
62	-63.0953	-61.3762	1.7191
63	-63.1586	-61.4088	1.7498
64	-63.3115	-61.3688	1.9427
65	-62.5499	-63.5485	-0.9986
66	-69.0272	-61.3666	7.6606
67	-63.2662	-61.3733	1.8929
68	-62.1784	-61.5883	0.5901
69	-66.8158	-61.3685	5.4473
70	-68.2889	-61.3669	6.9220
71	-63.0209	-61.3781	1.6428
72	-63.1896	-61.4501	1.7395
73	-62.4734	-61.5525	0.9209
74	-72.0509	-61.3666	10.6843

Figure 4.2 A comparison of log-likelihood when samples of bankrupt (+) and surviving (O) groups of companies are increased by 1 for each of 74 unclassified companies.

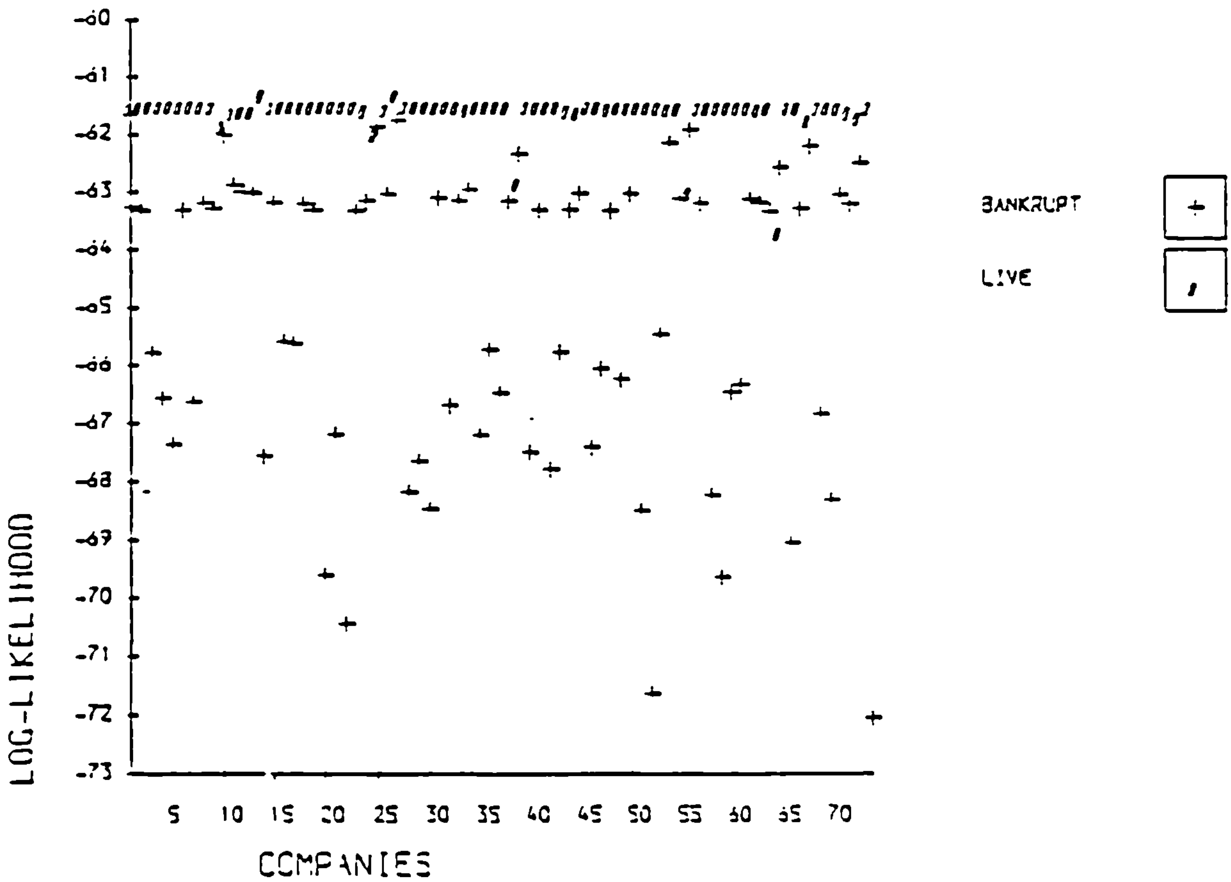
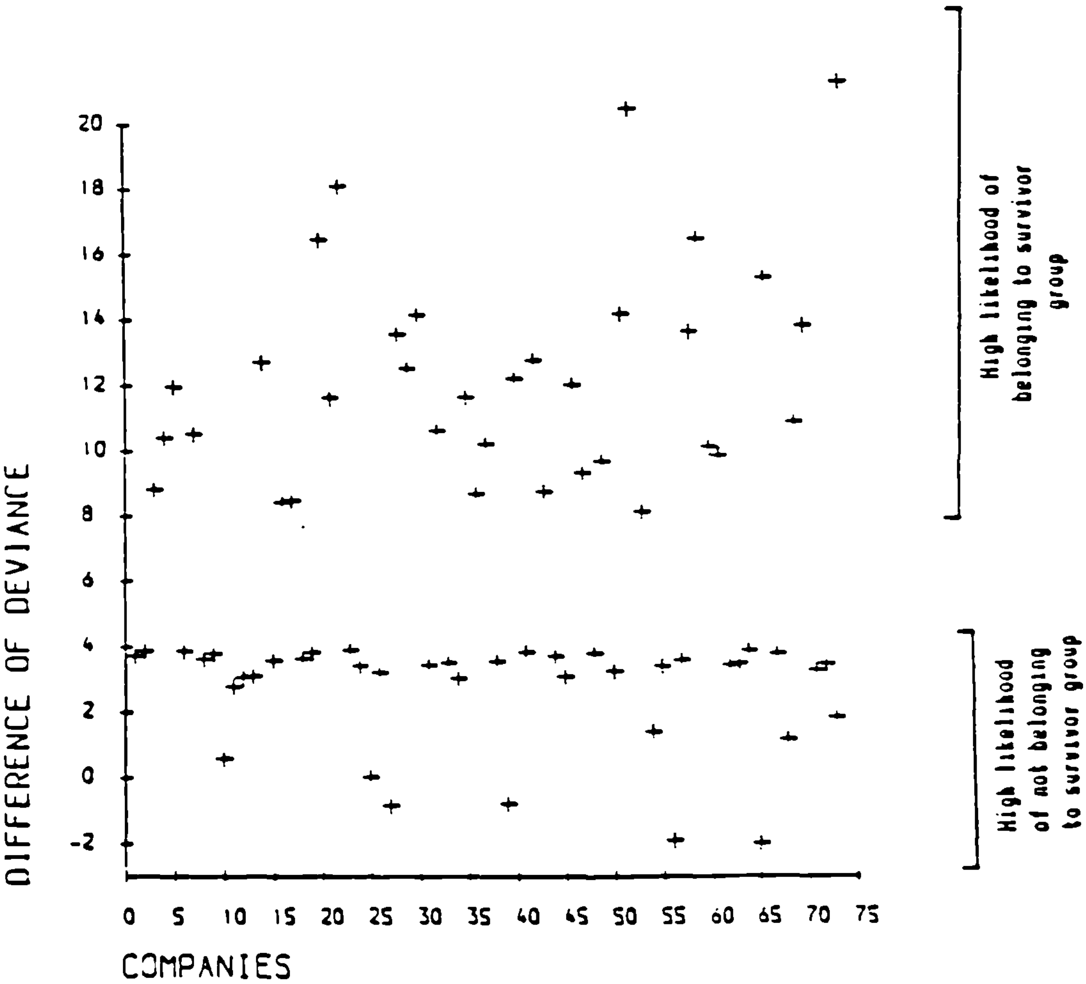


Figure 4.3 The differences between the deviance when the 74 unclassified companies are added separately to each of the two known groups.



A run of discriminant analysis was made *after* reclassifing the "merged" and "other" companies, where the sample of bankrupt companies was increased by 40, and the sample of surviving companies was increased by 34 (which we will name case b). The results in Table 4.7 show that the percentage correctly classified was 87.2% and type I error was 13.5% and type II error 12.8%. But, in the case of a quadratic discriminant function, we found that the percent of cases correctly classified was 67.6% and type I error was 29.5% and type II error 32.9%. Stepwise discriminant analysis showed that only 13 variables out of the 23 were significant. There were in descending order, Funds flow/current liabilities (FF/C.LIB), Current liabilities/total assets (C.LIB/TA), Sales/total assets (S/TA), Net income/net worth (NI/NW), Total assets/net worth (TA/NW), Retained earnings/total assets (RE/TA), Current assets/current liabilities (CA/C.LIB), EBIT/sales (EBIT/S), Funds flow/sales (FF/S), Funds flow/total assets (FF/TA), Current assets/total assets (CA/TA), Quick assets/current liabilities (QA/C.LIB) and Current assets/sales (CA/S).

Basically, the Survival Model method for reclassification in this case gave a higher percentage of correctly classified than when the unknown companies were left out from the analysis or when included with the bankrupts. Also, there was a lower type I error, as we have seen in Section 4.5. Therefore the Survival Model method is found with this data to be a more powerful method for reclassifying the unknown cases of "merged" and "other" companies into failures and non-failures.

We may conclude that the discriminant model, *after* reclassification of the "merged" and "other" companies using Survival Model method (case b), provides a useful discriminant function, because it gives a higher percentage correctly classified (87.2%) and lower type I error (13.5%) and type II error (12.8%) than obtained previously (see earlier sections of this chapter).

The results of applying a linear discriminant analysis *after* reclassification using the Survival Model method for 1, 2, 3, 4 and 5 years prior to failure are given in Table 4.8(a) and(b). It can be seen from Table 4.8(a) and (b) that, applying linear discriminant analysis to prior years, increased the type I error and type II error. But a randomly-censored stratified sampling technique (see pages 100-102) decreased the type I error in all cases -see Table 4.8(a) and (b).

Table 4.7 Results of linear discriminnt analysis *after* reclassification of "merged" and "other" companies by using bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group.

Actual group	Method	prior prob.	Predicted group membership		percentage correctly classified
			1	2	
non-survivor (1)	stepwise (Wilks)	equal (1:1)	86.5%	13.5%	87.2%
survivor (2)			12.8%	87.2%	

Table 4.8 Results of linear discriminant analysis *after* reclassification of merged and "other" companies (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) for five years prior to failure by using equal prior probabilities.

(a) - survivor data: one year prior to censoring

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error (%)	33.3	62.3	51.9	62.5	70.7
Type II error (%)	27.5	30.5	32.4	26	31.2
Correctly classified (%)	71.7	65.5	65.1	69.2	64.4
No. of cases					
non-survivor	61	57	52	51	41
survivor	384	363	343	335	330

(b) -survivor data: randomly-censored stratified samples

Type of model	No. of years prior to failure				
	1	2	3	4	5
Type I error (1) (%)	28.4 (3.5)	41.6 (9)	38.9 (6.5)	45.5 (6.7)	49.2 (5.1)
Type II error (1) (%)	30.6 (3.4)	29.6 (2.5)	31.4 (3.7)	24 (3.6)	32.3 (5.9)
Correctly (1) classified (%)	68.8 (2.6)	68.8 (1.6)	67.5 (3)	72.1 (2.4)	65.3 (4.5)
No. of cases					
non-survivor	61	57	52	51	41
survivor (1)	354	321	291	263	229

(1) All the values are averages based on 32 runs.

Figures in parenthesis are standard deviations.

Conclusion

In addition to providing benchmark results for comparison with applications of logit and survival models (see Chapter 5 and 6), it has been demonstrated in this chapter how (i) the use of a stratified sampling procedure to generate unequal samples of surviving companies which reflect the proportion of non-survivors to survivors but which nevertheless are matched to the lifetimes of the non-survivors and (ii) the use of a Survival Model likelihood method to reclassify those non-survivors which ceased trading without going into bankruptcy each has had the effect of decreasing type I and type II errors using the standard approach to failure prediction, i.e. discriminant analysis.

CHAPTER FIVE

A PROBABILISTIC MODEL OF FAILURE

(LOGISTIC DISCRIMINATION)

5.1 Introduction

In this chapter the econometric methodology of conditional logit analysis is used to avoid some fairly well known problems associated with multivariate discriminant analysis (MDA). This approach has been the most popular technique for failure studies using vectors of predictors. Some of the problems associated with MDA are :

(1) Certain statistical assumptions regarding the distributional properties of the predictors need to be satisfied for the validity of MDA. These assumptions are that predictors should have multivariate normal distribution and that their variance-covariance matrices should be the same for all groups e.g. two groups of non-surviving and surviving companies.

(2) The output of the application of an MDA model is a score which has little intuitive interpretation, since it is basically an ordinal ranking (discriminatory) device. For decision problems such that a misclassification structure is an inadequate description of the payoff partition, the score is not directly relevant, i.e. the payoff partition will be inadequate if it is not feasible to define a utility function over the two types of classification errors. Any economic decision problem would typically require a richer state partition. If however, prior probabilistics of the two groups are specified, then it

is possible to derive posterior probabilities of failure. But, this Bayesian revision process will be invalid or lead to poor approximations unless the assumption of normality is satisfied.

(3) There are also certain problems related to the matching procedures which have typically been used in MDA. For example, non-surviving and surviving companies are matched according to criteria such as size and type of industry. It is by no means obvious what is really gained or lost by different matching procedures, including no matching at all. At the very least, it would seem to be more beneficial actually to include variables as predictors rather than to use them for matching purposes (Ohlson, 1980).

The use of conditional logit analysis, on the other hand, essentially avoids the above difficulties. The fundamental estimation problem involved here can be reduced simply to the following statement: given that a company belongs to some pre-specified population, what is the probability that the company fails within some pre-specified time period? No assumptions need to be made regarding prior probabilities of failure and the analysis does not restrict the explanatory variables to any specific distributional form.

5.2 Choice of Predictor Variables

Given K variables X_1, X_2, ...,X_k observed on n individual companies where each belongs to one of two populations such as non-surviving or surviving groups of companies, if for each company there is also a Y variable which we wish to predict using a linear combination of X values, then we seek constants a_0, a_1,.....,a_k which minimise the residual sum of squares $\sum (Y-\hat{Y})^2$ where $\hat{Y} = a_0 + a_1X_1 + a_2X_2 + ... + a_kX_k$ is the linear predictor of Y. The most important question in multiple regression is to decide which of the predictor variables provide useful information about Y, and which may be safely omitted. The aim is to obtain a predictor using a small number of variables because this is simpler to interpret and often leads to more reliable prediction. The decision to include or exclude a particular variable X_i is equivalent to testing the hypothesis that the corresponding coefficient a_i is zero. This may be done in one of two equivalent ways:-

(i) let R be the residual sum of squares when k regressor variables are used and let $R_{(i)}$ be the residual sum of squares when X_i is excluded, but the other (k-1) variables remain. Clearly $R_{(i)} \geq R$, since the latter makes use of more information. Let

$$F = (R_{(i)} - R) / \{ R/(n-k-1)\}.$$

Then we conclude that X_i does provide significantly useful information about Y, in addition to that provided by the other X's, if F is large.

(ii) one may calculate the standard error, S_i of $\hat{a}_i$, the estimate of a_i. This will indicate that a_i is significantly different from zero if the t-value $| t | = |\hat{a}_i \div S_i|$ is large compared to the critical value of the t-distribution with n-k-1 degrees of freedom (Morrison, 1976).

5.2.1 Stepwise Regression

The problem with the above approach is that the significance, or otherwise, of a particular variable depends critically on which other variables are also being considered at the same time. This often leads to some complexity and confusion when there are more than three or more variables. A useful automatic technique is stepwise regression which involves entering variables into the regression one at a time, then considering whether any of the other variables currently in the equation should be dropped. At each stage one selects the variable to enter which reduces the residual sum of squares the most, and the variable to drop is the one whose omission increases the residual sum squares the least, i.e. has the smallest t-value. This process continues iteratively until the situation stabilises with the F-values being less (or greater) than specified critical values F-to-enter (or F-to-remove). In this way a large number of predictor variables may be reduced automatically to a smaller set. Larger critical values lead to a smaller set of regressors remaining. Details of this technique is discussed in Draper and Smith (1981).

5.3 Generalised Linear Models

Classical multiple regression analysis as described above is based on the assumptions that

(a) the expected value of Y, μ say, can be represented as a linear combination of the regressors.

$$\mu = a_0 + \sum_i a_i X_i \quad , \text{ and}$$

(b) that Y is distributed normally about the mean with a constant variance.

Generalised linear models (Mccullagh and Nelder 1983) allow for

(a)$'$ some function f, called the link function which represents the relationship between the mean of the ith observation and its linear predictor, such that

$$f(\mu) = a_0 + \sum a_i X_i \quad , \text{ and}$$

(b)$'$ the response Y may be distributed according to one of several kinds of probability distributions, including the normal.

Once these have been specified the parameters are estimated by the method of maximum likelihood. In the case of the identity link, $f(\mu) = \mu$, and a normal response these assumptions correspond exactly to (i) and (ii) of multiple regression.

5.3.1 Prediction of Binary Variables - Logistic and Logit Models

If the response variable is binary, i.e. Y takes the values 0 (surviving companies) or 1 (non-surviving companies), this corresponds to a Bernoulli random variable with

$E(Y) = p$, say, and

$Pr(Y = 1) = p$,

$Pr(Y = 0) = 1 - p$

which may be written as

$$Pr(Y) = p^{Y} (1 - p)^{1-Y} \quad , \qquad Y = 0, 1.$$

This may be thought of as a Binomial distribution Bin(n, p) with $n = 1$.

For given values of regressor variables, $X_1, X_2, \ldots, X_K$ we aim to predict p, the probability of failure for a company within a specified period of time. Because p is constrained to the range [0, 1] and for easy interpretation it is usually transformed and the most commonly used are the probit and the logit transformations. As they generally give similar results we use the logit as our link function since it is simpler.

In logit analysis, the outcome or response variable is a binary variable which records the event of surviving ("success") or non-surviving ("failure"). The predicted proportion of successes, s/n where s is the number of successes and n is the total number of cases (successes plus failures), follows the logistic model

$$p = \frac{e^{\theta}}{1 + e^{\theta}} = [\ 1 + \exp(-\theta)]^{-1}, \qquad \text{.........(5.1)}$$

where $\theta = a_0 + \sum_{i=1}^{k} a_i X_i$, is a linear function of the predictors. This model is non-linear in θ and the probability p approaches 0 or 1. The linear predictor θ represents the incremental effects of the X's and as θ increases, there is an increase in the probability of failure.

When θ = 0 , p = 1/2 . Thus, the probability of failure is greater than 1/2 when the sign of θ is positive and less than 1/2 when the sign is negative (see Figure 5.1).

The logit link or transformation defined by

$$\log \left[\frac{p}{1 - p}\right] = \theta = a_0 + \sum_{i=1}^{k} a_i X_i \qquad \text{..........(5.2)}$$

maps the range [0, 1] of p to (- ∞ , ∞) for the logit function, becomes linear in the predictors and represents the log odds of a company failing. The odds of failure

$$\frac{p}{1 - p} = e^{\theta} = e^{a_0 + \sum_{i=1}^{k} a_i X_i}$$

is an appealing interpretation of this model.

Thus e^{a_i} is the change in odds of a company failing per unit increase in the predictor X_i. If a_i is positive then $e^{a_i} > 1$ which implies that the failure odds are increased, while the odds are decreased if a_i is negative since $e^{a_i} < 1$ and the odds will be unchanged if $a_i = 0$ since $e^{a_i} = 1$.

The parameters, a's, of the logistic model (5.1) are estimated by the maximum likelihood method and the procedure is iterative since the model is non-linear. The log likelihood function is given by (Ohlson, 1980)

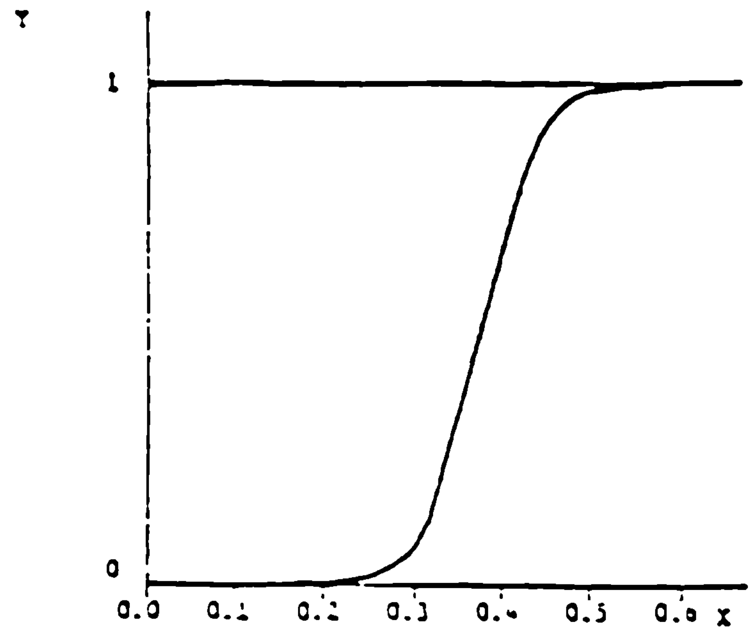
$$\log L(a_0, a_1, \ldots, a_k) = \sum_{i \in S1} \log\,[1 + \exp(-\theta)]^{-1} + \sum_{i \in S2} \log\,[1 - \{1 + \exp(-\theta)\}]^{-1} \quad \ldots(5.3)$$

which reflects the binary sample space consisting of non-surviving (S_1) and surviving (S_2) companies.

In model (5.1) the value of p, may be interpreted as the conditional probability of failure for a company with a given set of financial characteristics. Companies are predicted to fail if this probability exceeds the critical of 1/2, (Gentry et al 1985, Hamer 1983); they are predicted not to fail if $P < 1/2$. Two types of error are possible. Type I error is defined as predicting that a non-surviving company will survive and type II error is defined as predicting a surviving company will fail (Meryer and Pifer, 1970).

Here for a given set of regressor values, we estimate the probability that an individual company belongs to one of two population and in the context of company diagnosis, this seems a natural and useful procedure to employ. The technique will be used to classify companies as to whether they suffer failure or not, and to determine the extent to which the predictor variables affect the risk of company failure.

Figure 5.1 The logit cumulative distribution function



5.4 Applications

This section reports our empirical results for the stepwise reduction of variables and the probabilistic model techniques applied to the financial data base that was collected for this study.

The analysis is carried out on the same data sets; cases a(1 and 2) and b (Chapter 4); *before* and *after* reclassifying "merged" and "other" companies. The computer packages SAS and MINITAB are used for analysing the data.

5.4.1 Stepwise Reduction of Variables

Because 23 predictor variables is too large a set to study intelligently, an automatic procedure was used to reduce the number of predictors to the most significant or "best" subset and this achieved through stepwise regression of failure on the predictors. The following table shows the significant variables identified by the analyses for both cases a(1 and 2) and b, given in their order of importance for each case. The three, groups of variables are basically the same except for the differences in their order of selection . However, the ratio QA/TA appears in case a(2) as the most significant variable but does not feature in either case a(1) or case b. Also the ratio NW/S is selected for case a(1) and a(2) but not for case b.

Table 5.1 Stepwise selection of "best" subset results for cases a(1 and 2) and b *before* and *after* reclassifying "merged" and "other" companies.

a before reclassifying "merged" and "other"		b after reclassifying "merged" and "other" (non-survivor(bankrupt, and 40 others) and survivor(surviving and 34 others)groups)
(1) non-survivor (bankrupt) and survivor groups	(2) non-survivor (bankrupt,"merged" and "other")and survivor groups	
variable	variable	variable
NW/S	QA/TA	FF/C.LIB
FF/C.LIB	C.LIB/TA	C.LIB/TA
C.LIB/TA	FF/C.LIB	S/TA
CA/TA	EBIT/S	NI/NW
CA/C.LIB	RE/TA	TA/NW
NI/NW	CA/C.LIB	RE/TA
QA/C.LIB	S/TA	CA/C.LIB
RE/TA	NW/S	EBIT/S
EBIT/S	FF/TA	FF/S
FF/S	FF/S	FF/TA
FF/TA	TA/NW	CA/TA
CA/S	QA/C.LIB	QA/C.LIB
	NI/NW	CA/S

5.4.2 Application of the Generalised Linear Model

The procedure of generalised linear model with logit link (Section 5.3) is used here to predict company failure with the predictors being those significant variables which were identified by stepwise regression analysis (Table 5.1) for cases a(1 and 2) and b *before* and *after* reclassifying "merged" and "other" companies. The results are summarised in Tables 5.2 and 5.3 respectively.

The tables show regression coefficients of the stated predictor variables with the probability of failure p related linearly to the predictors in the form $\log\left[\frac{p}{1-p}\right] = \sum_{i=0}^{k} a_i X_i$, where the constant a_0 standing for the grand mean. A positive coefficient indicates increased probability of failure with increasing values of the variable concerned.

We infer from Table 5.2 for case a(1) that there are nine significant indicators of failure. High values of NI/NW, FF/S, CA/TA, FF/C.LIB and QA/C.LIB decrease the probability of failure while high values of FF/TA, EBIT/S, CA/C.LIB and C.LIB/TA increase that probability. In case a(2) that there are eleven financial ratios that significantly indicate company failure. High values of NI/NW, FF/S, S/TA, TA/NW, FF/C.LIB, RE/TA and QA/C.LIB lead to low incidence of failure, while high values of FF/TA, EBIT/S, CA/C.LIB and C.LIB/TA increase the incidence.

From Table 5.3 (case b) we note that there are eleven significant indicators of failure with probability of failure increasing as the values of FF/TA, EBIT/S, CA/C.LIB and C.LIB/TA increase, while this probability decreases as the values of NI/NW, FF/S, CA/TA, S/TA, TA/NW, FF/C.LIB and QA/C.LIB increase.

Table 5.2 Results of logit analysis for case a(1 and 2) *before* reclassifying "merged" and "other" companies.

a(1) non-survivor(bankrupt) and survivor groups				a(2) non-survivor(bankrupt, "merged" and "other" companies) and survivor groups			
variable	coeff-icient	chi-sq.	prob.	variable	coeff-icient	chi-sq.	prob.
NW/S	0.57	1.66	0.1981	QA/TA	-1.12	1.44	0.2302
FF/C.LIB	-5.07	23.61	0.0001	C.LIB/TA	3.21	16.41	0.0001
C.LIB/TA	4.10	23.75	0.0001	FF/C.LIB	-2.73	20.77	0.0001
CA/TA	-6.93	34.85	0.0001	EBIT/S	8.99	29.82	0.0001
CA/C.LIB	1.33	28.74	0.0001	RE/TA	-0.99	9.10	0.0026
NI/NW	-2.81	14.59	0.0001	CA/C.LIB	0.68	19.77	0.0001
QA/C.LIB	-1.23	12.09	0.0005	S/TA	-1.08	31.74	0.0001
RE/TA	-0.67	1.34	0.2479	NW/S	0.03	0.01	0.9245
EBIT/S	13.13	24.51	0.0001	FF/TA	11.54	36.28	0.0001
FF/S	-7.72	7.11	0.0077	FF/S	-8.45	21.70	0.0001
FF/TA	11.22	10.91	0.0010	TA/NW	-0.36	91.40	0.0001
CA/S	0.71	2.57	0.1091	QA/C.LIB	-0.91	7.73	0.0054
a_0	-2.04	7.16	0.0074	NI/NW	-3.44	34.90	0.0001
				a_0	-0.53	0.98	0.3215

Table 5.3 Results of logit analysis for case b *after* reclassifying "merged" and "other" companies.

variable	coefficient	chi-sq.	prob.
FF/C.LIB	-4.34	42.65	0.0001
C.LIB/TA	4.61	26.79	0.0001
S/TA	-1.85	28.33	0.0001
NI/NW	-2.19	16.38	0.0001
TA/NW	-0.34	13.49	0.0002
RE/TA	-0.49	1.80	0.1794
CA/C.LIB	1.05	31.41	0.0001
EBIT/S	6.54	20.65	0.0001
FF/S	-5.00	8.77	0.0031
FF/TA	11.87	36.02	0.0001
CA/TA	-1.91	5.73	0.0166
QA/C.LIB	-1.23	26.41	0.0001
CA/S	-0.81	2.72	0.0993
a_0	0.48	0.80	0.3697

5.4.3 Prediction of Failure

From the fitted model

$$\log \left[\frac{p}{1 - p}\right] = \sum_{i=0}^{k} a_i X_i$$

the estimated probability of failure for a given set of values of the predictors is

$$p = \frac{e^{\Sigma a_i X_i}}{(1 + e^{\Sigma a_i X_i})}$$

Table 5.4 shows overall correct classification, type I and type II errors for cases a(1 and 2) and b by using p = 0.5 as a critical value to classify the data. We conclude from the table that the logit model for case b is the best model which gives the highest percentage correctly classified (88.7%) cases and the lowest type I and type II errors (11.5 and 11.2 respectively).

Table 5.4 Overall correct classification, type I and type II errors for cases a(1 and 2) and b *before* and *after* reclassifying "merged" and "other" companies.

Type of model	case a		case b
	(1)	(2)	
Type I error	14.8%	16.2%	11.5%
Type II error	16.8%	17.9%	11.2%
percentage correctly classified	83.4%	82.6%	88.7%

Table 5.5 shows overall correct classification, type I and type II errors for cases a(1 and 2) and b for the linear discriminant analysis , quadratic discriminant analysis and logit analysis.

We infer from this table that the logit model provides a modest increase in the overall correct classification rate and a decrease in type I and type II errors for cases a(1 and 2) and b over the linear discriminant model but a substantial improvement over the quadratic discrimination. Since the purpose of a model is to identify companies that are likely to fail with reasonable accuracy, the improvement provided by the logistic model after reclassification using survival analysis are of some value.

Table 5.5 A comparison of linear discriminant analysis, quadratic discriminant analysis and logit analysis for cases a(1 and 2) and b *before* and *after* reclassifying "merged" and "other" companies.

		linear discriminant analysis	quadratic discriminant analysis	logit analysis
a(1)	Type I error	15.5%	38.1%	14.8%
	Type II error	18%	25.1%	16.8%
	percentage correctly classified	82.3%	74.1%	83.4%
a(2)	Type I error	19.5%	33.7%	16.2%
	Type II error	22.3%	34.1%	17.9%
	percentage correctly classified	78.2%	64.9%	82.6%
b	Type I error	13.5%	29.5%	11.5%
	Type II error	12.8%	32.9%	11.2%
	percentage correctly classified	87.2%	67.6%	88.7%

CHAPTER SIX

SURVIVAL MODELS

6.1 Introduction

Much of the empirical analysis in the issue of time series financial analysis, with particular respect to corporate failure has been concerned with discriminating between non-surviving and surviving companies, more recently with a view to obtaining parsimonious models of the characteristics of non-surviving companies from extensive data sets. Logit analysis has also been used to assess the likelihood of failure, and recursive partitioning has been used to model the stepwise procedures inherent in financial analysis when screening out the potential type one errors.

In this chapter we have taken a different approach, that of survival analysis, and we have attempted to model the covariates of survival, in an attempt to understand the structure of large cross-sectional accounting information sets. The analysis of survival data has received considerable attention in the last decade and comprehensive accounts are now available (Burridge, 1982). The principal ones being Mann, Schafter and Singpurwalla (1974), Barlow and Proscham (1975), Gross and Clark (1975) and Kalbfleisch and Prentice (1980). The last of these contains several recent developments including some of the material considered in this chapter.

Survival analysis is concerned with the analysis of a population where, for each individual or company, we observe either the time to failure or the time to censoring. For censored individuals, the time to failure is a random variable. Lifetime is denoted by T, and the probability of a company surviving to time t is given by

$$S(t) = Pr(T \geq t)$$

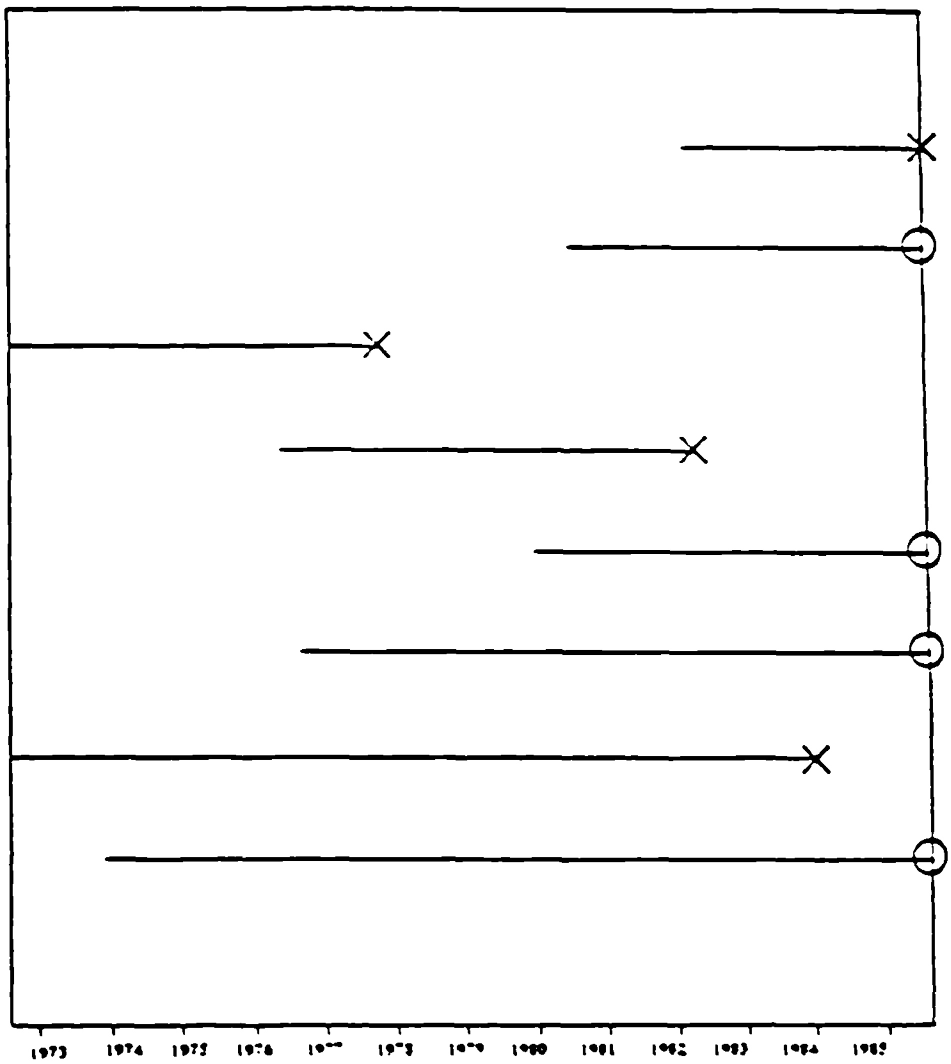
where S(t) is a survival function.

Most companies have staggered entry, so that they enter over a substantial time period. Each company's failure time is usually measured from its own date of entry. Figure 6.1 illustrates the calculation.

Within the companies area a typical question which arises relating to a failure, is: how is the progress of failure affected by the characteristics of a company such as its financial ratios? To answer such a question as this a study is often carried out which involves looking at the length of time companies survive from the beginning of the study until some event of interest (failure). This time is called the survival time, and survival analysis is the area of statistics used to model it.

Although most of the applications in the literature to which the methods of survival analysis have been applied are medical, the possible applications range from the industrial, such as the accelerated testing of rubber tyres under factory conditions (Davis 1985), to the social/economic of determining which factors are likely to affect a person's return to full time employment following a period of unemployment (Lancaster and Nickell 1980).

Figure 6.1 Eight companies with staggered entry, failed (x) or censored (o).



6.2 Survival Function and Hazard Function

We consider a population of individuals where for each individual we observe either the time to failure or the time to censoring. That is for the censored individuals we known only that the time to failure is greater than the censoring time. Let T be a non-negative random variable representing the lifetime of an individual from a homogeneous population. The probability of an individual surviving till time t is given by

$$S(t) = Pr(T \geq t) \qquad \text{............(6.1)}$$

called the survivor function (Lawless, 1982). From the definition of the cumulative distribution function F(t) of T,

$$S(t) = 1 - F(t) \qquad \text{............(6.2)}$$

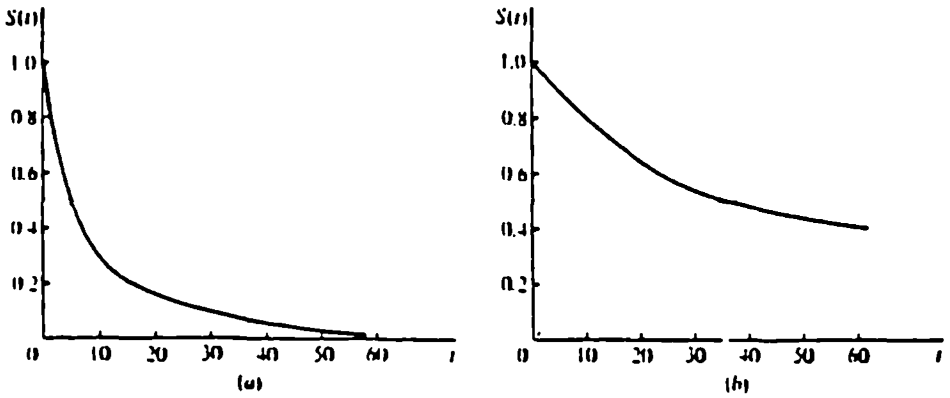
survivor function S(t) is a nonincreasing function of time t with the properties (Cox and Oakes, 1984),

$$S(t) = \begin{cases} 1 & , \text{ for } t = 0 \\ 0 & , \text{ for } t = \infty \end{cases}$$

that is, the probability of surviving at least at the time 0 is one and that of surviving an infinite time is zero.

The graph of S(t) is called the survival curve. A steep survival curve, such as the one in Figure 6.2(a), represents low survival rate or short survival time. A gradual or flat survival curve such as in Figure 6.2(b) represents high survival rate or longer survival.

Figure 6.2 Two examples of survival curves.



A fundamental concept in survival analysis is that of the hazard function h(t), which is defined as the conditional density function at time t given survival up to time t (Aitkin et al, 1989), i.e.

$$h(t) = \lim_{\Delta t \to 0} \frac{pr(\, t \le T < t + \Delta t \mid T \ge t \,)}{\Delta t} \qquad \text{......(6.3)}$$

The hazard function can also be defined in terms of the cumulative distribution function F(t) and the probability density function $f(t)$:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \qquad \text{...........(6.4)}$$

The hazard function specifies the instantaneous rate of death or failure at time t, given that the individual survives up to time t. In particular $h(t)\Delta t$ is the approximate probability of death in $[t, t + \Delta t)$, given survival up to t.

The hazard function may increase, decrease, remain constant, or indicate a more complicated process (Nelson, 1972). Figure 6.3 plots several kinds of hazard functions. For examples, $h_1(t)$ is an increasing hazard function where the rate of failure increases with time, $h_2(t)$ decreases with time, $h_3(t)$ is where the rate of failure is constant, $h_4(t)$ is called bathtub curve, it reflects the process of human life where the death rate declines, remains constant and then increases with age, and $h_5(t)$ describes a process such as corporate bankruptcy where the failure rate increases sharply after incorporation but then declines with survival time.

The density and survivor function can be obtained from hazard function as,

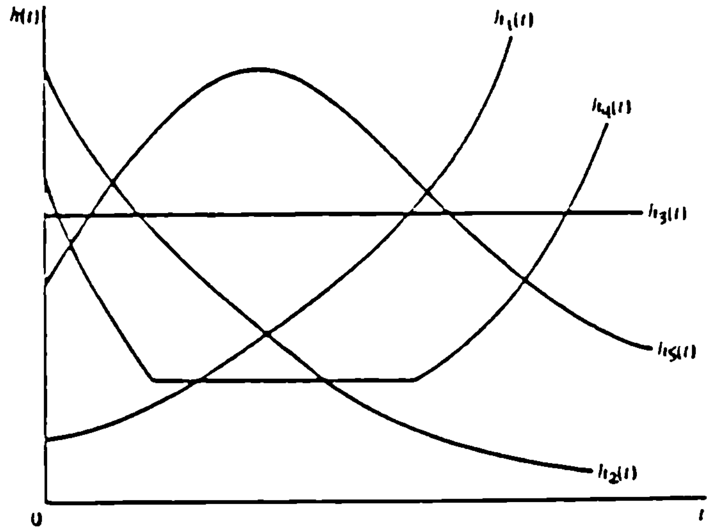
$$S(t) = \exp [- H(t)] \qquad \text{......(6.5)}$$

$$f(t) = h(t) \exp [- H(t)] \qquad \text{.......(6.6)}$$

where, H(t) is the cumulative hazard function given by

$$H(t) = \int_0^t h(X)\, dX = \int_0^t \frac{f(X)}{1 - F(X)}\, dX = -\log S(t).$$

Figure 6.3 Examples of the hazard function



6.3 Estimation of the Survival Function

The product-limit developed by Kaplan-Meier (1958) is used for estimating the survivor function. This method is applicable for any sample size, small, moderate, or large. However when the sample size is very large it may be convenient to group the survival times into intervals and perform a life-table analysis. The product-limit and life-table estimates of the survivor function are essentially the same. Many authors use the term life-table estimates for the product-limit estimates.The only difference is that the product-limit estimate is based on individual survival times while in the life-table method survival times are grouped into intervals (Lee, 1980). If there are no censored observations in a sample of size n the empirical survivor function is defined as

$$\hat{S}(t) = \frac{\text{number of observations} \geq t}{n} \quad , \quad t\geq 0 \qquad \ldots.(6.7)$$

This is a step function the decreases by 1/n just after each observed lifetime if all observation are distinct. More generally if there are d lifetimes equal to t the empirical survival function drops by d/n just past t. When dealing with censored data some modification of (6.7) is necessary since the number of lifetimes greater than or equal to t will not generally be known exactly. The modification of (6.7) described is called the Kaplan-Meier product-limit estimate of the survivor function. The estimate is defined as follows : suppose that there are observations on n individuals and that there are k ($k \leq n$) distinct times $t_1 < t_2 < \ldots.. < t_k$ at which deaths may occur.

There is no real loss of generality in assuming that these times are discrete, because the finite precision of measurement means that the values of survival time actually recorded can take only a finite (though possibly large) number of values (Aitkin et al, 1989). The possibility of there being more than one death at t_j is allowed, and we let d_j represent the number of deaths at t_j. In addition to the lifetime $t_1,\ldots\ldots,t_k$, there are also censoring times L_i for individuals whose lifetimes are not observed. The product-limit estimate of S(t) is defined as

$$\hat{S}(t) = \prod_{j:t_j<t} \frac{n_j - d_j}{n_j} \qquad \ldots\ldots\ldots\ldots(6.8)$$

where n_j is the number of individual at risk at t_j, that is the number of individuals alive and uncensored just prior to t_j (Lawless, 1983). The function (6.8) is a nonparametric maximum likelihood estimate in the family of all possible distribution (Kaplan and Meier, 1958).

6.4 Parametric Models

Usually there are many causes that lead to the failure of an individual company at a particular time. It is often difficult to isolate these causes and mathematically account for all of them. Therefore choosing a theoretical distribution to approximate survival data is as much an art as a science. Even though our interest in this section concerns the relationship between failure time and explanatory variables it is necessary to consider briefly failure time distribution for homogeneous populations. We will look at three theoretical distributions that have been widely used to describe failure time.

(1) Exponential model

The simplest and most important distribution in survival studies is the exponential distribution. In the late 1940's and early 1950's, researches chose the exponential distribution to describe the life pattern of electronic systems (Lee, 1980). The one parameter exponential distribution is obtained by taking the hazard function to be a constant, $h(t) = \lambda > 0$, over the range of lifetime T. This means that the exponential model has a constant hazard function which implies that the probability of death at time t is not dependent on the length of previous lifetime, i.e., the instantaneous probability of failure is the same no matter how long the item has already survived. A large λ indicates high risk and short survival while a small λ indicates low risk and long survival. When the survival time T follows the exponential distribution with a parameter λ, the probability density function $f(t)$ is defined as (This is well known distribution before 1966)

$$f(t) = \begin{cases} \lambda\ e^{-\lambda t} & , \quad t \geq 0,\ \lambda > 0 \\ 0 & , \quad t < 0 \end{cases} \qquad \text{.....(6.9)}$$

The cumulative distribution function is

$$F(t) = 1 - e^{-\lambda t} \qquad , \qquad t \geq 0 \qquad \text{......(6.10)}$$

and the survivor function is then

$$S(t) = e^{-\lambda t} \qquad , \qquad t \geq 0 \qquad \text{.......(6.11)}$$

so that, the hazard function is

$$h(t) = \lambda \qquad , \qquad t > 0 \qquad \text{........(6.12)}$$

The probability density function of Y = log T is then,

$$\text{Exp}\ (Y - \alpha - e^{Y - \alpha}) \qquad , \qquad -\infty < Y < \infty \quad \text{..(6.13)}$$

where

$\alpha = -\log \lambda$

(2) Weibull model

The Weibull distribution is a generalization of the exponential distribution. However, unlike the exponential distribution, it does not assume a constant hazard rate therefore has broader application. The Weibull distribution is characterized by two parameters, γ and λ that determine the shape of the distribution curve and its scale. The relationship between the value of γ and survival time can be seen as follows : when $\gamma = 1$, the hazard rate remains constant as time increases; this is the exponential case. The hazard rate increases when $\gamma > 1$ and decreases when $\gamma < 1$ as t increases. Thus the Weibull distribution may be used to model the survival distribution of a population with increasing, decreasing or constant risk (Aitkin et al, 1983).

The probability density function is

$$f(t) = \lambda\gamma(\lambda t)^{\gamma - 1}\exp[-(\lambda t)^{\gamma}] \quad , \quad t \geq 0,\ \gamma,\ \lambda > 0 \qquad \ldots\ldots\ldots(6.14)$$

The survivor function is, therefore

$$S(t) = \exp[-(\lambda t)^{\gamma}] \qquad \ldots\ldots\ldots(6.15)$$

and the hazard function, the ratio of equation (6.14) to (6.15) is

$$h(t) = \lambda\gamma(\lambda t)^{\gamma - 1} \qquad \text{.........(6.16)}$$

The probability density function of the log failure time Y (Y = log T) is (Lawless, 1982)

$$\sigma^{-1}\exp\left[\frac{Y - \alpha}{\sigma} - e^{(Y - \alpha)/\sigma} \right] \quad , \quad -\infty < Y < \infty \qquad \text{..(6.17)}$$

where $\sigma = \gamma^{-1}$ and $\alpha = -\log\lambda$

More simply we can write $Y = \alpha + \sigma w$, where w has the extreme value density function.

(3) <u>Log-logistic model</u>

The log-logistic has slightly heavier density in the tails, and is often used for survival data that is left- and right-censored. Here the probability density function and survivor function are given by

$$f(t) = \lambda\gamma(\lambda t)^{\gamma - 1} [1 + (\lambda t)^{\gamma}]^{-2} \qquad \text{......(6.18)}$$

and

$$S(t) = \frac{1}{[\ 1 + (\lambda t)^{\gamma}\ ]} \qquad ......(6.19)$$

and the hazard function is

$$h(t) = \frac{\lambda\gamma(\lambda t)^{\gamma - 1}}{[\ 1 + (\lambda t)^{\gamma}\ ]} \qquad ......(6.20)$$

Again, the model for $Y = \log T$ is of the form

$$Y = \alpha + \sigma w \qquad ......(6.21)$$

where w has logistic density $\dfrac{e^{w}}{(\ 1 + e^{w}\ )^{2}}$, $\lambda = e^{-\alpha}$ and $\gamma = \sigma^{-1}$

The three distributions were introduced above for modeling the survival time of a homogeneous population. Usually however there are explanatory variables upon which failure time may depend. It therefore becomes of interest to consider generalizations of these models to take account of concomitant information on the individual sample. An important method of handling heterogeneity in a population is through the inclusion of regressor variables in the model (Kalbfleisch and Prentice, 1980). It is common for data to involve regressor variables related to lifetime. For example lifetimes of industrial U.K. companies may depend on factors such as FF/TA, NI/NW, EBIT/S, QA/TA, FF/S, NW/S, S/TA, TA/NW, FF/C.LIB, RE/TA, CA/C.LIB and C.LIB/TA financial ratios. Suppose now that on each individual one or more further measurements are available, say on variables $X_1, X_2, \ldots\ldots, X_K$. The main problem considered in this chapter is that of assessing the relation between the distribution of failure time and the X's. For example the exponential distribution can be generalized to obtain a regression model by allowing the failure rate to be a function of the covariates X (Turnbull et al, 1974). The hazard at time t for an individual with covariates X may be written as

$$h(t;X) = \lambda(X)$$

Thus the hazard for a given X is a constant characterizing an exponential failure time distribution, but the failure rate depends on X. The $\lambda(X)$ function may be parameterized in many ways. If the effect of the components of X is only through a linear function, $X\beta$ we have

$$h(t;X) = \lambda\ c(X\beta)$$

where $\beta' = (\beta_1, \ \beta_2, \ldots\ldots, \beta_K)$ is a vector of regression parameters, λ is a constant and c is a specified functional form. The specific forms that have been used is c(X) = exp(X), which is the most natural form since it takes only positive values (Kalbfliesch and Prentice, 1980). Consider then the model with hazard function

$$h(t;X) = \lambda\ e^{X\beta'} \qquad \ldots\ldots\ldots(6.22)$$

The conditional density function of T given X is then

$$f(t;X) = \lambda\ e^{X\beta} \exp\ (\ -\ \lambda t\ e^{X\beta}\ )$$

$$= \exp\ (\ X\beta\ -\ te^{X\beta}\ ).$$

The model (6.22) specifies that the log failure rate is a linear function of the covariates X. In terms of the log survival time, Y = log T, the model (6.22) may be written as

$$Y = \alpha - x\beta + W \qquad \ldots\ldots\ldots(6.23)$$

where $\alpha = -\log\lambda$ and W has the extreme value distribution with the probability density function

$$\exp\ (W - e^{W}) \qquad , \qquad -\infty < W < \infty$$

6.4.1 Maximum Likelihood Estimation

Let t_1, t_2,......,t_n, t_{n+1},......,t_{n+m} be the survival times of n + m individuals the last m of which are censored. Let X_{ij}, for $i = 1,\ldots\ldots, n + m$, and $j = 0, 1, \ldots\ldots, k$ be the corresponding values of explanatory variables with $X_{i0} = 1$. The survival time has density function $f(t)$, distribution function $F(t)$ and hazard function

$$h(t) = \frac{f(t)}{S(t)}, \qquad \text{where } S(t) = 1 - F(t).$$

The hazard function is assumed to involve the explanatory variables through a log-linear model (Aitkin and Clayto, 1980) as

$$h(t_i) = \lambda(t_i)\ \exp\ (\beta' X_i).$$

Thus the density function $f(t)$ is assumed to be of the form

$$f(t) = \lambda(t) \exp\left(\beta' X - H(t)\, e^{\beta' X}\right),$$

and hence $S(t) = \exp\left(- H(t)\, e^{\beta' X}\right)$, where

$$H(t) = \int_0^t \lambda(u)\, d(u).$$

Let δ be an indicator variable taking the value 1 for uncensored, and 0 for censored observations. Under the usual assumption that the censoring mechanism is independent of the explanatory variables, the likelihood function (also see Chapter 4 Section 4.6.2) is

$$L = \prod_{i=1}^{n+m} [\, f(t_i)\,]^{\delta_i} \, [\, S(t)\,]^{1 - \delta_i}$$

$$= \prod_i \, [\, \lambda(t_i) \exp(\beta' X_i)\,]^{\delta_i} \exp\left(- H(t)\, e^{\beta' X}\right) \quad \ldots(6.24)$$

The unknown parameters involved are estimated by using maximum likelihood and the resulting equations are solved by the Newton-Raphson method to obtain the maximum likelihood (M.L.E) estimates $\hat{\beta}$. If there is little or no censoring, initial estimates can be obtained using the least squares ignoring censoring (Lawless, 1982).

6.4.2 Residual Analysis

The examination of residuals from a fitted model is an important tool for checking the assumption of the model (Nelson, 1973). Suppose Y_i is a response variable and X_i is an associated vector of regressor variables. The distribution of Y_i given X_i, is specified except for a vector β of unknown parameters and we assume that the model can be represented in terms of quantities

$$e_i = g_i(Y_i, \beta, X_i) \qquad ....(6.25)$$

that are independently identically distribution (i.i.d.) and whose distribution is known. If $\hat{\beta}$ is the M.L.E. of β, determined from data (Y_i, X_i), then the residuals $\hat{e}_i$ are defined by (Lawless, 1982).

$$\hat{e}_i = g_i(Y_i, \hat{\beta} , X_i) \qquad ....(6.26)$$

These residuals are often considered as behaving approximately like a random sample of size n from the distribution of e_i. For example suppose that the distribution of T_i given X_i were exponential and has survivor function

$$S(t|X) = \exp (- te^{-X\beta}).$$

Since the quantities $e_i = (t_i e^{-X\beta})$ are i.i.d. with standard exponential distribution, residuals could be defined by

$$\hat{e}_i = (t_i e^{-X_i\hat{\beta}}) \quad , \qquad i = 1, \ldots\ldots, n \qquad .....(6.27)$$

When the data are censored, modifications are necessary :if one observes a censoring time rather than a lifetime the corresponding residual is censored as well. One approach in this situation is to treat the observed residuals, both censored and uncensored as a censored sample from the distribution of e_i. The product-limit estimator or empirical hazard function can then be calculated from the residuals for an estimate of the underlying survivor function of the e_i's. Plots of this estimate can be used to assess the underlying distribution (Lawless, 1982).

Suppose that the cumulative hazard function of T_i given X_i is $H(t_i|X_i)$, since $S(t_i|X_i) = \exp[-H(t_i|X_i)]$, $i = 1, \ldots, k$ are i.i.d. random variables uniformly distribution on (0, 1) the $H(t_i|X_i)$'s are i.i.d. standard exponential random variables. If residuals for uncensored observations $t_1, \ldots, t_k$ are defined by

$$\hat{e}_i = \hat{H}(t_i|X_i) \qquad \ldots\ldots(6.28)$$

where $\hat{H}(t_i|X_i)$ involves the M.L.E.'s of unknown parameters, then as first approximation $\hat{e}_1, \ldots, \hat{e}_k$ can be treated as a random sample from the standard exponential distribution (Lagakos, 1980). Residuals for censored observation if defined as in (6.28) can be treated as censored standard exponential observations. One can then form a product-limit survivor function estimate from the set of censored and uncensored residuals. If the residual is based on (6.28), a plot of $-\log[\hat{S}(e_i)]$ versus. e_i should give roughly a straight line with slope 1, when the model is adequate.

6.5 Nonparametric Model (Proportional Hazards Model)

The multiple regression method is a conventional technique for investigating the relationship between survival time and possible predictor variables (Prentice and Gloeckler, 1978). Let X_1, ..., X_p be p possible predictor variables. For the ith company the observed values of the p variables are denoted by X_{1i}, ..., X_{pi}. In the multiple regression approach the survival time of the ith company t_i is the dependent variable, depending on the values of the p independent variables. We are interested in identifying a relationship of t_i or a function of t_i, say $g(t_i)$ and (X_{1i}, ..., X_{pi}) that may be expressed in a regression form as

$$t_i = f_1(X_{1i}, \ldots\ldots, X_{pi})$$

or

$$g(t_i) = f_2(X_{1i}, \ldots, X_{pi})$$

Regression models proposed for survival distributions generally involve the assumption of proportional hazard function. One such model introduced by Cox (1972) is a general nonparametric model appropriate for the analysis of survival data with and without censoring. The proportional hazards model, however is nonparametric in the sense that it involves an unspecified function in the form of an arbitrary base-line hazard function (Miller et al, 1981).

In estimating the hazard rate for a company it is also important to recognise the role played by financial ratios as well as the history of company failure (as was stated in Section 6.4). These predictor variables, often called covariates, represent inherent differences among the companies in a study. Regression models allow us to incorporate this additional structure into the estimation of the hazard rate. These models may defined via the equation (Lawless, 1983)

$$h(t,X) = g(\beta, X)\ \lambda_0(t) \qquad \text{.........(6.29)}$$

where $h(t,X)$ is the hazard function for a company with regressor variables X, $g(\beta,X)$ expresses the relationship between X and the regression parameters β and $\lambda_0(t)$ is an unknown function giving the hazard function for the standard set of conditions, X = 0.

This section is based upon the Cox (1972) proportional hazards regression model. The model is formulated in terms of the effects of the covariates upon failure (hazard) rates rather than upon times to failures (i.e. the effect of the covariates is to act multiplicatively on the hazard function) (Anderson et al, 1985). Cox (1972) suggested a regression model for the failure time t of an company, where one or more further measurements are available on variables $X_1, \ldots, X_K$. For t continuous, the proportional hazards model is specified by the hazard function (Ingram and Kleinman, 1989),

$$h(t,X) = \lambda_0(t)\ \exp(\beta' X) \qquad \text{..........(6.30)}$$

or written in log-linear forms

$$\text{Log}\left[\frac{h(t,X)}{\lambda_0(t)} \right] = \beta' X \qquad \text{..........(6.31)}$$

where X is a vector of covariates (financial ratios), β is a set of parameters to be determined and $\lambda_0(t)$ is an unspecified function of t often known as the underlying hazard. One of the major advantages of this model is that an estimate of β can be obtained using the method of likelihood function (Cox, 1972) or the method of partial likelihood (Cox, 1975) which does not depend on the particular form of the underlying hazard. There will be a loss of efficiency in estimating β by not specifying $\lambda_0(t)$ if known, however it has been shown that this is often small (Cox and Oakes, 1984). The loss of efficiency can be offset by a gain in robustness for the estimator of β due to not having to specify the underlying hazard. This model is now widely used.

Cox (1975) suggested that the properties of the partial likelihood estimator $\hat{\beta}$ would be similar to those of a maximum likelihood estimator, i.e. would be asymptotically normal, unbiased and have variance given by the inverse of the information matrix. His heuristic arguments have been strengthened by the work of Andersen and Gill (1982) who derived the properties of the partial likelihood estimate, having reformulated the problem within the theory of counting processes.

6.5.1 Model Assumptions

Although the underlying hazard function is unspecified, the basic model, in which the covariates do not depend on time, does make the following assumptions (see Schoenfeld, 1980 and Moreau and et al, 1985) :

(1) Proportional Hazards Assumption.

The first assumption is the multiplicative relationship between the underlying hazard function and the log-linear function of covariates. Thus the ratio of the hazard functions for two individuals with time independent covariate X_1 and X_2 given by

$$\frac{h(t,X_1)}{h(t,X_2)} = \exp\,[\beta'(X_1 - X_2)]$$

does not depend upon time. The assumption can be relaxed by specifying a particular parametric time dependency for some of the covariates. For example (Cox, 1972)

$$h(t,X) = \lambda_0(t)\ e^{(\beta' X + \gamma X_1 t)}$$

For a covariate X_1.

(2) Log Linear Assumption.

The second assumption of the model is the log-linear effect of the covariates upon the hazard function. Rewriting equation 6.30 we see that the log hazard function is proportional to the linear term $\beta' X$

$$\log h(\ t,\ X\ ) = \log \lambda_0(t) + \beta' X$$

6.5.2 Maximum Likelihood Estimates

For all companies in our study, one or more measurements are available on the p explanatory variables X_1, X_2, ...,X_p. For continuously distributed failure times T may be discrete or continuous (Cox, 1972). In our study, T is assumed to be continuous, the possibility of tied failure times can be ignored, and the hazard function is given by

$$h(t,X) = \lambda_0(t) \exp(\beta' X),$$

where β is a row vector of p unknown parameters. For the ith company the values of the covariate

$$X = (X_1, X_2, \ldots\ldots, X_p)' \text{ are } X_{ij} = (X_{1i}, X_{2i}, \ldots\ldots, X_{pi})',$$

where p is the number of explanatory variables. The parameters β are then estimated by maximising the likelihood function. No information can contributed about β by time intervals in which no failures occur, because the component $\lambda_0(t)$ might conceivably be identically equal to zero in such intervals.

Suppose that $t_{(1)} < t_{(2)} < \ldots\ldots < t_{(k)}$ represent k distinct times to failure among n observed survival times. For a particular failure at time $t_{(i)}$, the risk set, $\mathbb{R}(t_{(i)})$ is the subset of all those at risk when the ith failure occurs i.e. $\mathbb{R}(t_{(i)})$ consists of all companies whose survival times are at least $t_{(i)}$. Then for the particular failure at time $t_{(i)}$ conditional on the risk set $\mathbb{R}(t_{(i)})$ the probability that the failure of each company is as observed is given by (Crapp and Stevenson, 1987)

$$\frac{\exp(\beta' X_{(i)})}{\sum_{\ell \in \mathbb{R}(t_{(i)})} \exp(\beta' X_{(\ell)})} \qquad \ldots\ldots\ldots\ldots(6.32)$$

Multiplying these probabilities together for each of the k failure times gives the likelihood function

$$L(\beta) = \prod_{i=1}^{k} \frac{\exp(\beta' X_{(i)})}{\sum_{\ell \in R(t_{(i)})} \exp(\beta' X_{(\ell)})} . \qquad \ldots\ldots(6.33)$$

The likelihood is a conditional likelihood function since each of the probabilities in equation (6.32) are conditional probabilities, conditional on the risk set at a particular failure time $t_{(i)}$. Doubts have been raised as to whether equation (6.33) is indeed a likelihood function (see Kalbfleisch and Prentic, 1973). These doubts were refuted by Cox (1975) who showed that equation (6.33) produces inferences similar to ordinary likelihood procedures. The equation (6.33) does not depend on $\lambda_0(t)$. When all probabilities for each failure item are considered the conditional log-likelihood function is given by

$$\text{Log } L(\beta) = \sum_{i=1}^{k} \beta' X_{(i)} - \sum_{i=1}^{k} \log \left[\sum_{\ell \in R(t_{(i)})} \exp(\beta' X_{(\ell)}) \right]. \qquad \ldots(6.34)$$

We now obtain maximum likelihood estimate $\hat{\beta}$ of β from equation (6.34), using the Newton-Raphson method of iteration.

When there are ties among the failure times, Breslow (1974) suggests an approach to the estimation of β and $\lambda_0(t)$, which is different from those of Kalbfleisch and Prentice's (1973) and Cox (1972). The underlying survival distribution is assumed continuous, having constant hazard $h_i = \exp(\alpha_i)$ between each pair $(t_{(i)}, t_{(i+1)})$ of distinct failure times. All censored observations that occur in the interval $(t_{(i)}, t_{(i+1)})$ are assumed to have occurred at $t_{(i)}$. Then the likelihood proposed by Breslow (1974) is

$$L(\beta) = \prod_{i=1}^{k} [\exp(\beta' Z_i) / \{ \sum_{\ell \in \mathbb{R}(t_{(i)})} \exp(\beta' X_{(\ell)}) \}^{d_i}] \quad \ldots(6.35)$$

or the log-likelihood is

$$\text{Log } L(\beta) = \sum_{i=1}^{k} [\beta Z_i - d_i \log \sum_{\ell \in \mathbb{R}(t_{(i)})} \exp(\beta' X_{(\ell)})] \quad \ldots.(6.36)$$

where k is the number of distinct failure times ,d_i is the number of deaths at t_i and Z_i is the vector sum of covariates of the d_i individuals. When there are no ties $d_i = 1$ for i = 1, ...,k. Maximization of the likelihood function yields estimators of β.

6.5.3 A Stepwise Regression Procedure for the Selection of Variables

Our main interest here is to identify important prognostic factors. In other words we wish to identify from the p independent variables a subset of variables that relate significantly to the hazard and consequently the length of survival of the company. Recall that in a standard multiple regression problem (as we mentioned in Chapter 5) this can be achieved by using a stepwise regression method that ranks the variables in order of relative importance. From the ranking and the significance test for each variable we can select the most significant variables related to the dependent variables. Because of the close analogy between the standard multiple regression and the following equations (Lee, 1980)

$$\log \frac{h(t,X)}{\lambda_0(t)} = \beta' X_i$$

if we let $Y_i = \log \left[\frac{h(t,X)}{\lambda_0(t)} \right]$, then the above equation is simply

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots\ldots\ldots + \beta_p X_{pi} \qquad \ldots\ldots\ldots(6.37)$$

a stepwise regression can also be applied to equation (6.37).

In estimating $\beta_1, \ldots\ldots, \beta_p$ in Section 6.5.2 a stepwise procedure may be used to rank the variables. In a forward stepwise (or step-up) procedure, the independent variables are entered in the regression equation one at a time until the regression is satisfactory (Krall et al., 1975). The order of insertion is determined by using for example the maximum log-likelihood value, Log $L(\hat{\beta})$ as a measure of the importance of variables not yet in the regression equation. Using the maximum log-likelihood value as a measure it selects as the first variable to enter the regression equation, the variable say $X_{(1)}$, whose maximum log-likelihood is the largest. Let $\log L(\hat{\beta}_i)$, $i = 1,\ldots,p$ be the maximum log-likelihood value obtained from fitting only the ith prognostic variable. Then $X_{(1)}$ is the first variable to enter the regression if

$$\text{Log } L(\hat{\beta}_{(1)}) = \underset{i}{\text{Max}} \; [\text{Log } L(\hat{\beta}_i)]$$

Now there are (P - 1) prognostic variables not yet fitted. The maximum log-likelihood value log $L(\hat{\beta}_{(1)}, \hat{\beta}_{(i)})$ is computed for each of the (p - 1) independent variables and the one that gives the largest Log L value is the next variable to enter the regression equation. The procedure continues to fit one additional independent variable at a time until the regression is satisfactory. At every step a likelihood ratio test (Johson and Johnson, 1980) is performed to determine if the last variable entered adds significantly to the variables already selected. At the first step there is only one variable in the regression equation i.e.

$$\text{Log } \frac{h_i(t,X)}{\lambda_0(t)} = \hat{\beta}_{(1)} X_{(1i)}$$

where $X_{(1i)}$ the most important single variable related to hazard could be any one $X_1, \ldots\ldots, X_p$. To test the significance of $X_{(1)}$ we test the hypothesis $Ho: \beta_{(1)} = 0$. For this we treat

$$\chi^2 = \frac{[U(\beta_{(1)})]^2}{I(\beta_{(1)})} \qquad \ldots\ldots\ldots\ldots\ldots(6.38)$$

where U is the vector of first derivatives of the log-likelihood (equation 6.34 from Section 6.5.2) evaluated at regression parameters equal to zero and I is the corresponding matrix of second derivatives as chi-square distribution with one degree of freedom.

In this procedure the first variable selected is the most important single variable in predicting survival time, the second variable entered is the second most important etc. the process thus provides a successive selection and ranking of the independent variables according to their relative importance.

6.5.4 Residual Analysis

One of the most useful methods of assessing models that have been fitted to the data is by examining residuals (Aalen, 1989). The use of residuals in parametric models was discussed in Section 6.4.2 and similar procedures can be followed here. The simplest way to define residuals for the model (6.30) is to use (6.28). The residual corresponding to an uncensored lifetime is then (Lawless, 1982)

$$\hat{e}_i = \hat{H}\,(t_i|X_i) = \hat{H}_0(t_i)\,e^{X_i\hat{\beta}} = [\,-\log \hat{S}_0(t_i)]\,e^{X_i\hat{\beta}} \qquad \text{........(6.39)}$$

where $H_0(t)$ is the baseline cumulative hazard function and $\hat{H}_0 = -\log \hat{S}_0(t)$ is an estimate of it. The estimate $\hat{\beta}$ is the M.L.E. since quantities $H(t_i|X_i)$ are independent and have standard exponential distributions, the $\hat{e}$'s if there is no censoring should look roughly like a random sample from the standard exponential distribution. When there are censored observations the approach discussed in Section 6.4.2 can be used. For example residual $\hat{e}_i$ for censoring times t_i if defined as in (6.39) can be treated as censored standard·exponential observations. One can then form a product-limit survivor function estimate from set of censored and uncensored residuals. The resulting estimate $\hat{S}(e)$ should be consonant with an underlying standard exponential distribution, for example a plot of log $\hat{S}(e)$ versus e should be roughly linear with slope -1.

6.5.5 Estimation of the Survivor Function S(t;X)

Under the Cox proportional hazards model (Miller et al, 1981),

$$S(t;X) = \exp\left(-e^{\beta' X}\int_0^t \lambda_0(u)\,du\right)$$

$$= \exp\left(-e^{\beta' X} H_0(t)\right) = S_0(t)^{e^{\beta' X}} \qquad ....(6.40)$$

where $S_0(t) = e^{-H_0(t)}$ = an arbitrary survivor function .
To estimate S(t;X), we substitute $\hat{\beta}$ for β but how do we estimate $H_0(t)$ or $S_0(t)$? suppose now that data are available from the extended model (6.40) and consider the calculation of the non-parametric maximum likelihood estimate of $S_0(t)$. In doing this an argument analogous to that used in obtaining the Kaplan-Meier

estimate (Section 6.3) is employed. As before let $t_{(1)},\ldots\ldots,t_{(k)}$ be the distinct failure times let D_i be the set of labels associated with individual failing at $t_{(i)}$ and C_i be the set of labels associated with individuals censored in $[t_{(i)}, t_{(i+1)})$ $(i = 0,\ldots,k)$ where $t_{(0)} = 0$ and $t_{(k+1)} = \infty$. The censoring times in the interval $[t_{(i)}, t_{(i+1)})$ are t_l where l ranges over C_i. The contribution to the likelihood of an individual with covariate X who fails at $t_{(i)}$ is, under independent censorship,

$$S_0(t_{(i)})^{\exp(\beta X)} - S_0(t_{(i)} + 0)^{\exp(\beta X)}$$

and the contribution of a censored observation at time t is

$$S_0(t + 0)^{\exp(\beta X)} .$$

The likelihood function can then be written (Kalbfleisch and Prentice, 1980)

$$L = \prod_{i=0}^{k} \{ \prod_{l\in D_i} [S_0(t_{(i)})^{\exp(\beta X_l)} - S_0(t_{(i)} + 0)^{\exp(\beta X_l)}] \prod_{l\in C_i} S_0(t_l+0)^{\exp(\beta X_l)} \} \quad \ldots(6.41)$$

where D_0 is empty.

As with Kaplan-Meier estimate that L is maximized by taking $S_0(t) = S_0(t_{(i)}+0)$ for $t_{(i)} < t \le t_{(i+1)}$ and allowing probability mass to fall only at the observed failure time $t_{(1)}, \ldots\ldots, t_{(k)}$. These observations lead to the consideration of a discrete model with hazard contribution $1 - \alpha_j$ at $t_{(j)}$ $(j = 1, \ldots\ldots, k)$ thus we take

$$S_0(t_{(i)}) = S_0(t_{(i-1)}+0) = \prod_{j=0}^{i-1} \alpha_j \quad , \; i=1, \ldots\ldots, k$$

where $\alpha_0 = 1$.

On substitution in (6.40) and rearranging terms we obtain

$$\prod_{i=1}^{k} \Big[\prod_{j \in D_i} (1 - \alpha_i^{\exp(\beta X_j)}) \prod_{l \in \mathbb{R}(t_{(i)})-D_i} \alpha_i^{\exp(\beta X_l)} \Big] \qquad \ldots\ldots(6.42)$$

as the likelihood function to be maximized . We might take $\beta = \hat{\beta}$ as estimated from the likelihood function and then maximize (6.42) with respect to $\alpha_1, \ldots\ldots, \alpha_k$. Differentiating the logarithm of (6.42) with respect to α_i gives the maximum likelihood estimate of α_i as a solution to

$$\sum_{j \in D_i} \frac{\exp(\hat{\beta} X_j)}{1 - \hat{\alpha}_i^{\exp(\hat{\beta} X_j)}} = \sum_{l \in \mathbb{R}(t_{(i)})} \exp(\hat{\beta} X_l) \qquad \ldots\ldots\ldots(6.43)$$

if only a single failure occurs at $t_{(i)}$, (6.43) can be solved directly for $\hat{\alpha}_i$ to give

$$\hat{\alpha}_i = \left(1 - \frac{\exp(\hat{\beta} X_{(i)})}{\sum_{l \in \mathbb{R}(t_{(i)})} \exp(\hat{\beta} X_l)} \right)^{\exp(-\hat{\beta} X_{(i)})}$$

otherwise an iterative solution is required, a suitable initial value for the iteration is α_{i0} where

$$\log \alpha_{i0} = \frac{-d_i}{\sum_{l \in \mathbb{R}(t_{(i)})} \exp(\hat{\beta} X_l)}$$

which is obtained by substituting

$$\hat{\alpha}_i^{\exp(\hat{\beta} X_j)} = \exp(e^{\hat{\beta} X_j} \log \hat{\alpha}_i)$$

the maximum likelihood estimate of the baseline survivor function is then

$$\hat{S}_0(t) = \prod_{i | t_{(i)} < t} \hat{\alpha}_i \qquad \text{........(6.44)}$$

which is a step function like the Kaplan-Meier estimate with discontinuities at each observed failure $t_{(i)}$. When the covariate $X=0$ for all individuals sampled (6.44) reduces to Kaplan-Meier estimate. The estimated survivor function for covariate value X is

$$\hat{S}(t;X) = \prod_{i | t_{(i)} < t} \hat{\alpha}_i^{\exp(\hat{\beta} X)} \qquad \text{..........(6.45)}$$

6.6 Application of Methodology

In this section we shall apply the methodology of survival models to U.K. industrial companies. Failure is defined (as in Chapter 2, Section 2.2) as the phenomenon where the company exits from the industrial list due to implied pressures of financial distress. In other words, a failed company changes its current organisational form. The probability of failure, therefore is the probability that a company will leave the industrial register.

Most empirical studies that have examined the use of financial ratios as a means of predicting corporate failure assume that an organisation's financial ratios capture the influences of managerial policy, industrial and economic factors which are specific to an organisation's financial condition. In our study we ensure that all the variables which are thought to be potentially relevant as predictors of financial distress have been included in the initial set. The main objectives in this section is to assess which regressor variables are significantly related to survival time and to find the best fitted model for the data.

The computer packages SAS and MINITAB were used to carry out survival analysis by using the same cases a(1 and 2) and b in Chapter 4, Sections 4.5 and 4.6.2 (i.e. a(1 and 2) *before* reclassification of the "merged" and "other" companies and b *after* reclassification the "merged" and "other" companies).

6.6.1 <u>The Parametric Models</u>

Weibull, exponential and log-logistic regression models were fitted to the data. In these cases the regression variables (covariates) X act additively on Y (Y = log T) or multiplicatively on T. We use the 12 variables NW/S, FF/C.LIB, C.LIB/TA, CA/TA, CA/C.LIB, NI/NW, QA/C.LIB, RE/TA, EBIT/S, FF/S, FF/TA and CA/S obtained from stepwise regression discussed in Chapter 5, Section 5.2.1 for case a(1) (*before* reclassifying "merged" and "other") and the 13 variables QA/TA, C.LIB/TA, FF/C.LIB, EBIT/S, RE/TA, CA/C.LIB, S/TA, NW/S, FF/TA, FF/S, TA/NW, QA/C.LIB, and NI/NW for case a(2) (where the "merged and "other" are added to the bankrupt group). And the 13 variables FF/C.LIB, C.LIB/TA, S/TA, NI/NW, TA/NW, RE/TA, CA/C.LIB, EBIT/S, FF/S, FF/TA, CA/TA, QA/C.LIB and CA/S we used for case b (*after* reclassifying "merged" and "other" companies).

The results for case a(1) analysis are summarised in Table 6.1.

Table 6.1 Asymptotic likelihood inference for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using Weibull, exponential and log-logistic regression models .

Model	**Loglike-lihood**	**vari-ables**	**D.F.**	**Coeff-icient**	**S.E.**	**Chi-Sq.**	**P-value**
		intercept	1	4.06	0.93	20.22	0.000
		FF/TA	1	-18.66	8.20	5.18	0.023
		NI/NW	1	2.37	0.90	6.98	0.008
		EBIT/S	1	-3.47	2.59	1.79	0.181
		FF/S	1	-2.33	4.84	0.23	0.629
		CA/TA	1	4.47	2.18	4.22	0.040
Weibull	-60.37	NW/S	1	0.93	0.54	2.98	0.084
		FF/C.LIB	1	10.00	3.22	9.63	0.002
		RE/TA	1	-0.67	0.72	0.87	0.350
		CA/C.LIB	1	-0.55	0.74	0.55	0.458
		QA/C.LIB	1	1.28	0.65	3.92	0.048
		C.LIB/TA	1	-0.32	1.60	0.04	0.844
		CA/S	1	-2.33	0.74	9.90	0.002
		scale	1	0.52	0.10		
		intercept	1	3.97	1.68	5.56	0.018
		FF/TA	1	-33.02	14.05	5.52	0.019
		NI/NW	1	4.16	1.49	7.76	0.005
		EBIT/S	1	-5.91	5.15	1.32	0.251
		FF/S	1	-2.31	9.70	0.06	0.812
		CA/TA	1	8.17	3.72	4.82	0.028
Exponential	-65.60	NW/S	1	1.84	1.04	3.15	0.076
		FF/C.LIB	1	17.02	5.19	10.76	0.001
		RE/TA	1	-1.23	1.36	0.82	0.364
		CA/C.LIB	1	-1.16	1.37	0.71	0.399
		QA/C.LIB	1	2.36	1.19	3.90	0.048
		C.LIB/TA	1	-0.72	2.96	0.06	0.809
		CA/S	1	-4.19	1.21	12.03	0.001
		scale	0	1	0		
		intercept	1	3.82	1.03	13.77	0.0001
		FF/TA	1	-19.06	8.89	4.60	0.032
		NI/NW	1	2.70	1.01	7.13	0.008
		EBIT/S	1	-2.96	3.11	0.90	0.343
		FF/S	1	-3.90	5.29	0.54	0.461
		CA/TA	1	4.68	2.19	4.56	0.033
Log-logistic	-61.73	NW/S	1	1.15	0.62	3.41	0.065
		FF/C.LIB	1	10.45	3.43	9.27	0.002
		RE/TA	1	-0.65	0.77	0.70	0.402
		CA/C.LIB	1	-0.55	0.78	0.49	0.482
		QA/C.LIB	1	1.25	0.67	3.46	0.063
		C.LIB/TA	1	-0.05	1.86	0.00	0.978
		CA/S	1	-2.64	0.87	9.10	0.003
		scale	1	0.48	0.10		

From Table 6.1, it is clear that the Weibull regression model is to some extent preferable to the exponential and log-logistic regression models on account of the larger maximised log-likelihood (-60.37). There is as well a strong (significant) prognostic effect of the variables FF/TA, NI/NW, CA/TA, FF/C.LIB, QA/C.LIB and CA/S on survival time. The regression coefficients indicate the relationship between the covariates and survival time. A positive coefficient increases the value of survival time while a negative coefficient has the reverse interpretation. Also, this analysis indicates that there is no apparent dependence of survival time of U.K. industrial companies on the financial ratios EBIT/S, FF/S, RE/TA, CA/C.LIB and C.LIB/TA with NW/S being of marginal significance. The asymptotic $\chi^2_{(1)}$ statistics given in Table 6.1 are formed for the ith component as

$$\{\hat{\beta}_i \ / \ (\text{estimated standard error of } \hat{\beta}_i)\ \}^2.$$

The survival function S(t) is estimated ($\hat{S}(t)$) from the data by using the Kaplan-Meier method (see Section 6.3) and is plotted in Figure 6.4. It represents a very high survival rate or longer survival times. Also, the hazard function h(t) is useful in modeling survival time data. In many instances information is available as to how the failure rate will change with the amount of time on test, as can be seen from Figure 6.5. The Figure shows that h(t) increases, after an initial drop, it reaches its peak at approximately 70 months then decreases. In other words, the peak of company deaths occurs at 70 months.

Figure 6.4 Kaplan-Meier survival distribution function estimate for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group)

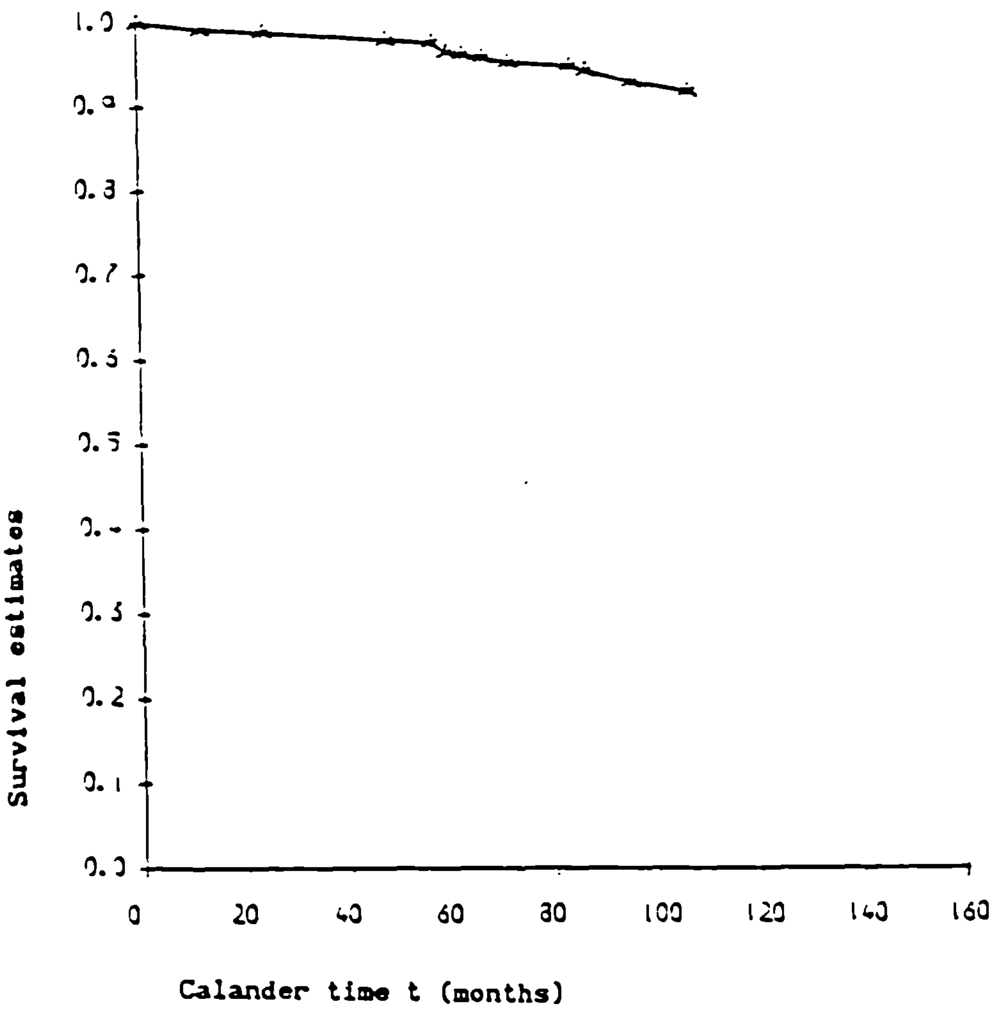
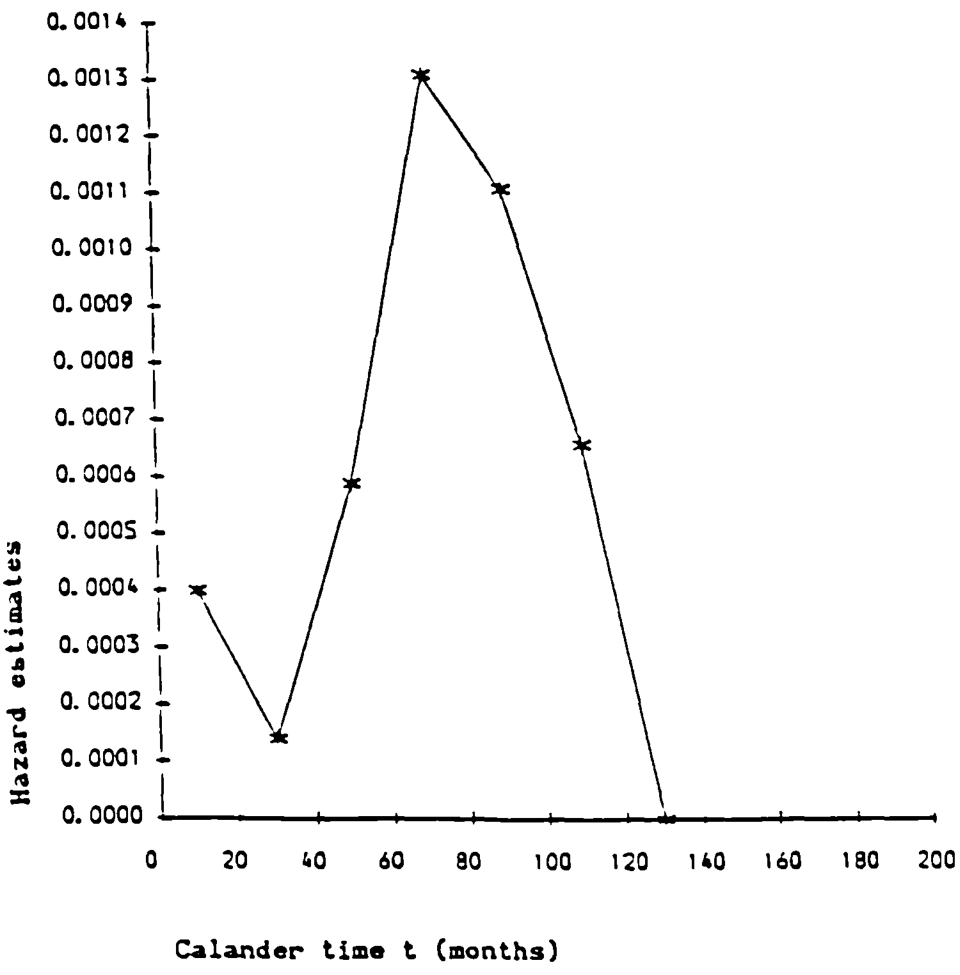


Figure 6.5 The hazard function estimate for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group)



Residual plots are used here to help assess the goodness of fit of the models. We plot the $-\log(\hat{S}(e))$ against ordered residuals , where $\hat{S}(e)$ is the product-limit estimate of the survivor function for the residuals e_i's, for all the three models assessed in Table 6.1. Figure 6.6 shows the plot of the points ($e_{(i)}$, $-\log(\hat{S}(e))$). The points lie roughly along straight line with approximately unit slope. It will be convenient to plot ($\log(e_i)$, $\log (-\log(\hat{S}(e)))$), shown in Figure 6.7, on the log scale or iterated log scale as this may give improvement to indicate the form of departure from the distribution assumption if this is incorrect. Again the points lie approximately on a straight line. a disadvantage of both the above plots is that for large e, when the number of observations on which $\hat{S}(e)$ is based becomes small, the variability of $\hat{S}(e)$ becomes very great and it is difficult to decide whether a real failure of the probability model is occurring. Similarly the variability of $\log(\hat{S}(e))$ becomes very great when e is small. Since $\hat{S}(e)$ is essentially an estimated binomial probability for each e, it can be "variance stabilized" using the arc sine transformation (Aitkin et al, 1989). Thus we plot $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$, as shown in Figure 6.8, in addition to the plots in Figures 6.6 and 6.7. The plot conforms closely to the straight line Y = X. Therefore the Weibull regression model appears to fit adequately. Figures 6.9, 6.10 and 6.11 show some evidence of curvature and these plots do not appear to suggest that the exponential model is appropriate.

Figures 6.12, 6.13 and 6.14, show the residual plots for log-logistics regression model. Again the plots are not close to straight lines. Therefore the log-logistic regression model also does not appear adequate.

Figure 6.6 Residual plot for Weibull regression model for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group)

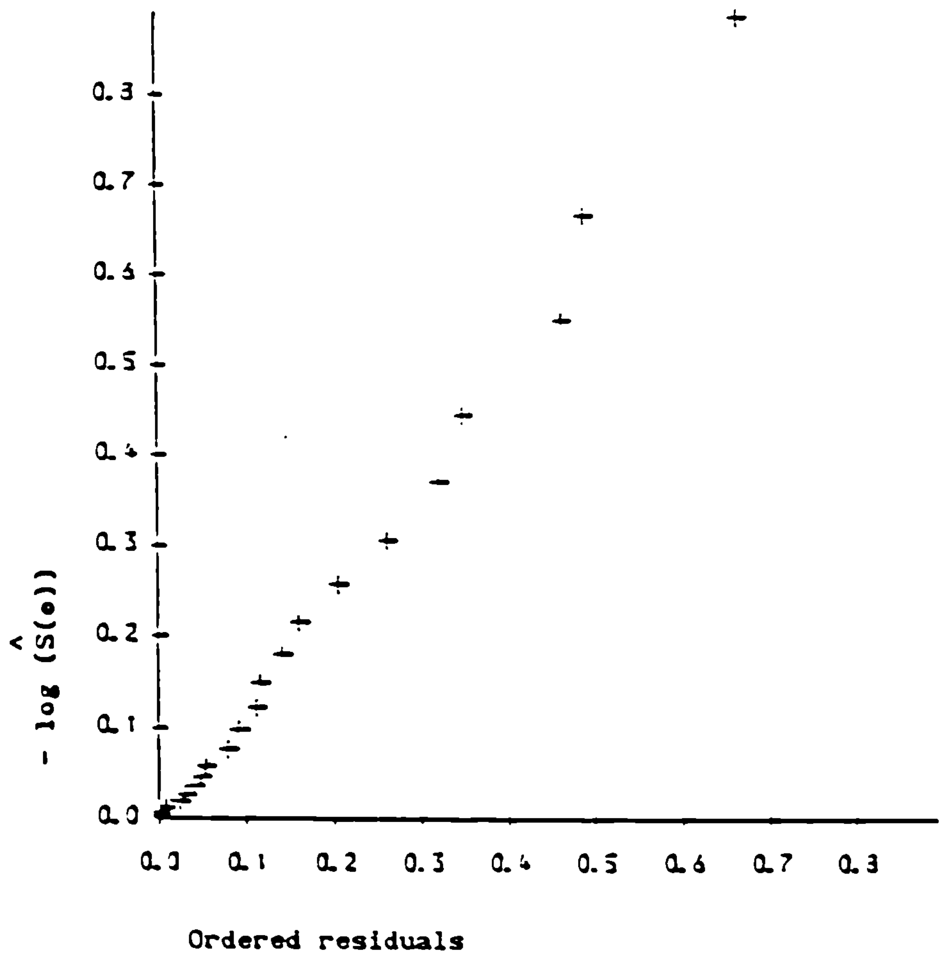


Figure 6.7 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using Weibull regression model.

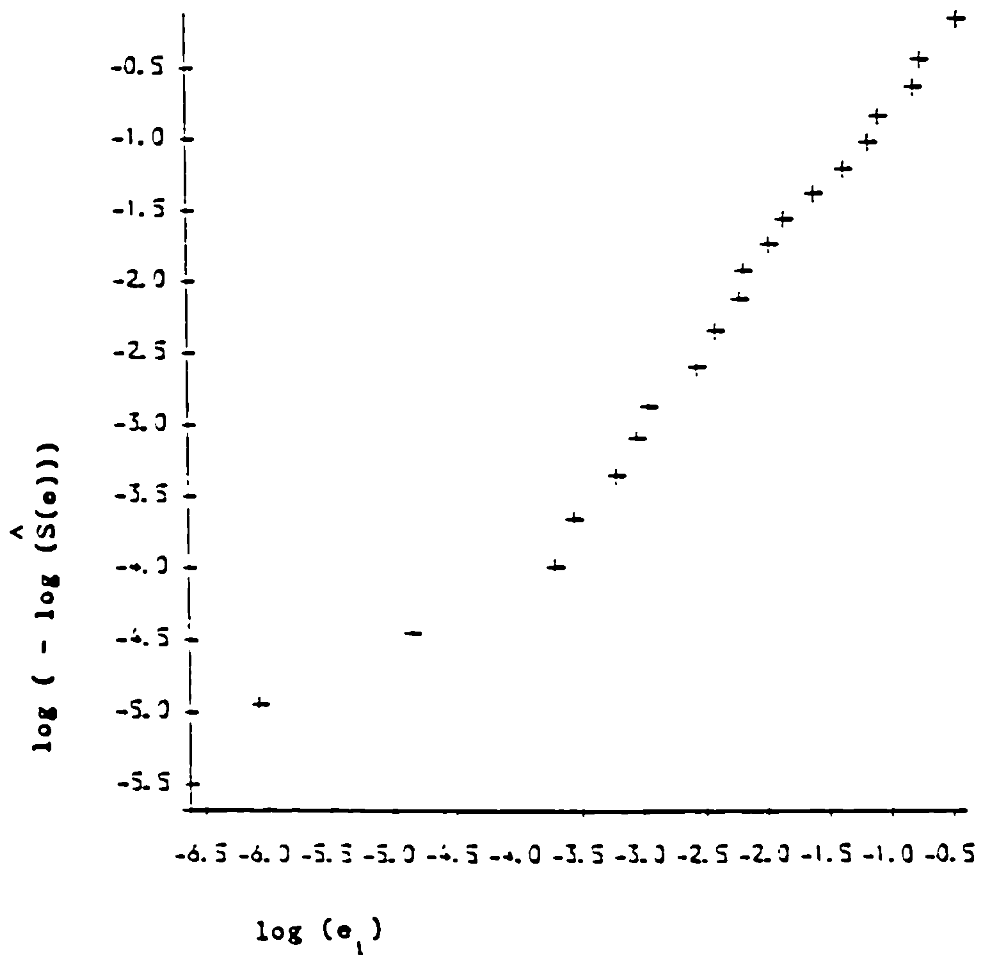


Figure 6.8 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using Weibull regression model.

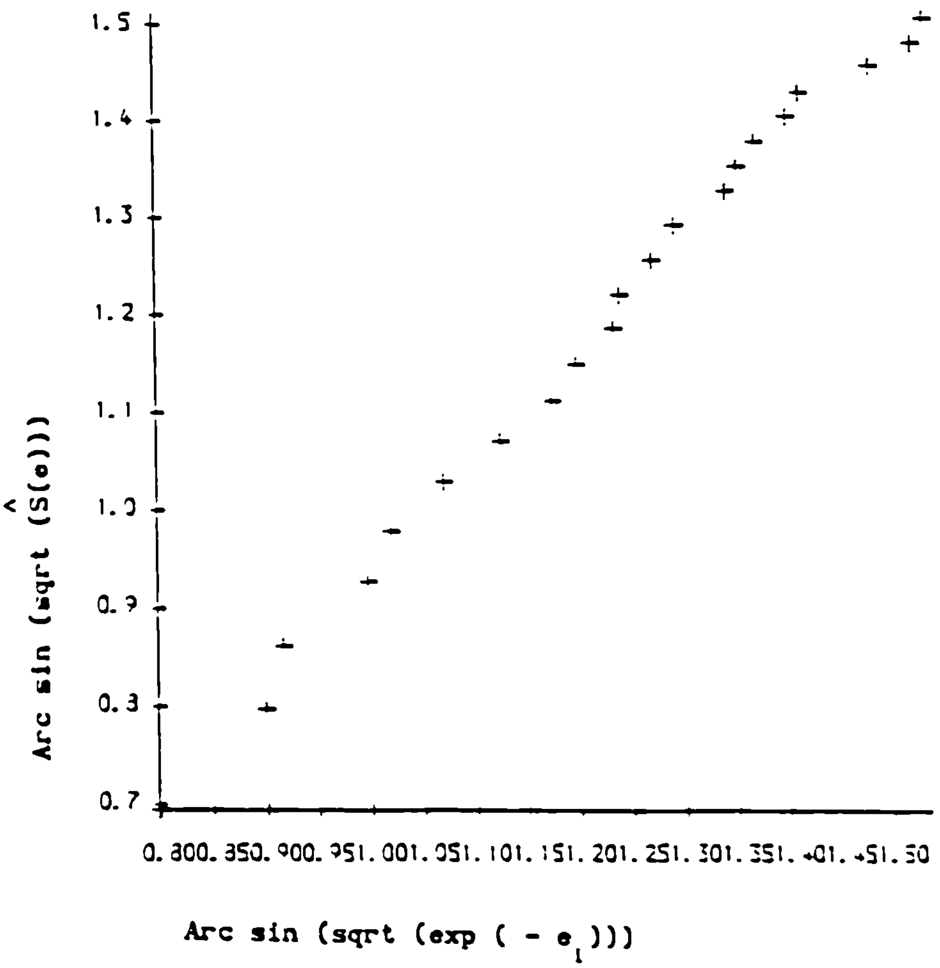


Figure 6.9 Residual plot for exponential regression model for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

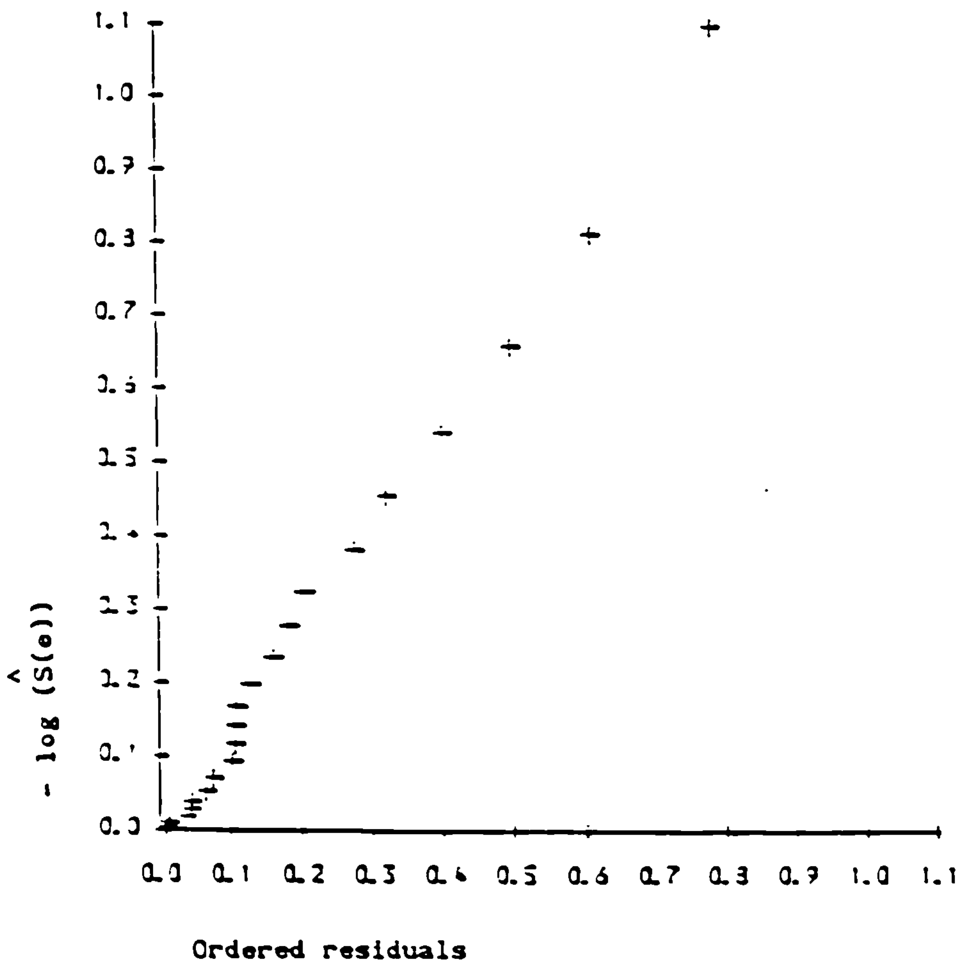


Figure 6.10 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using exponential regression model.

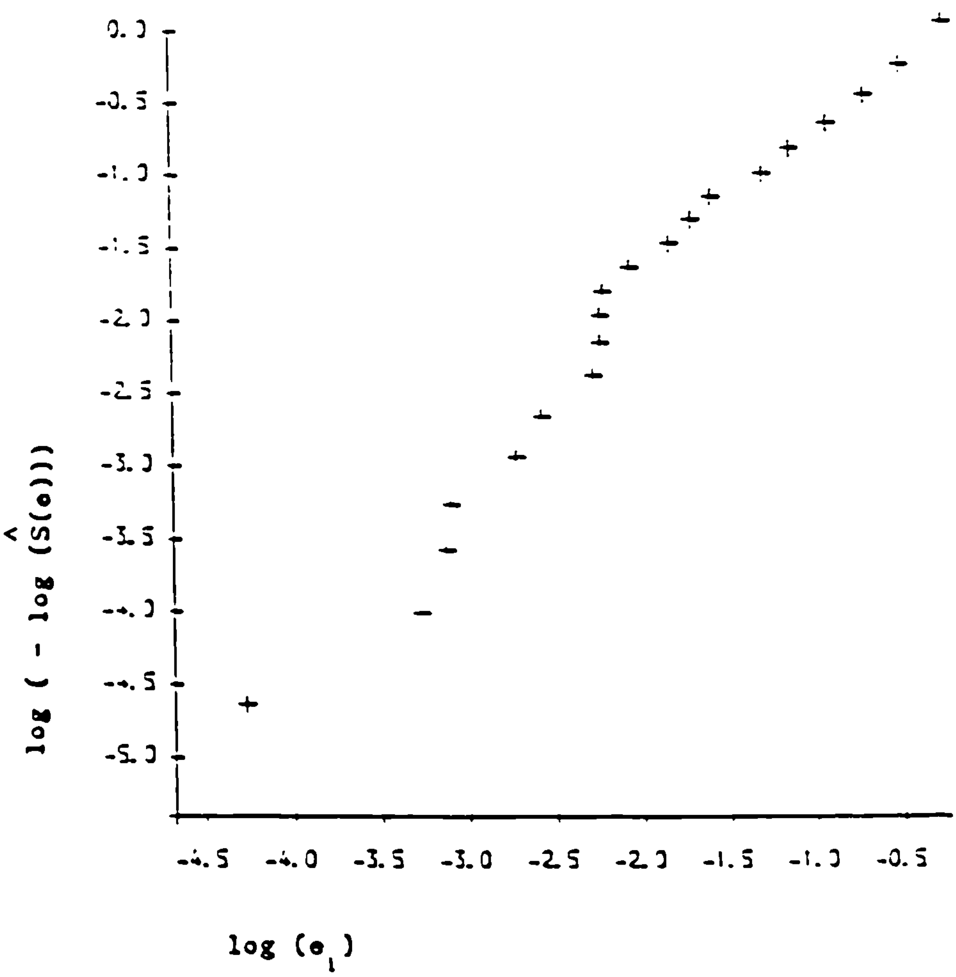


Figure 6.11 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using exponential regression model.

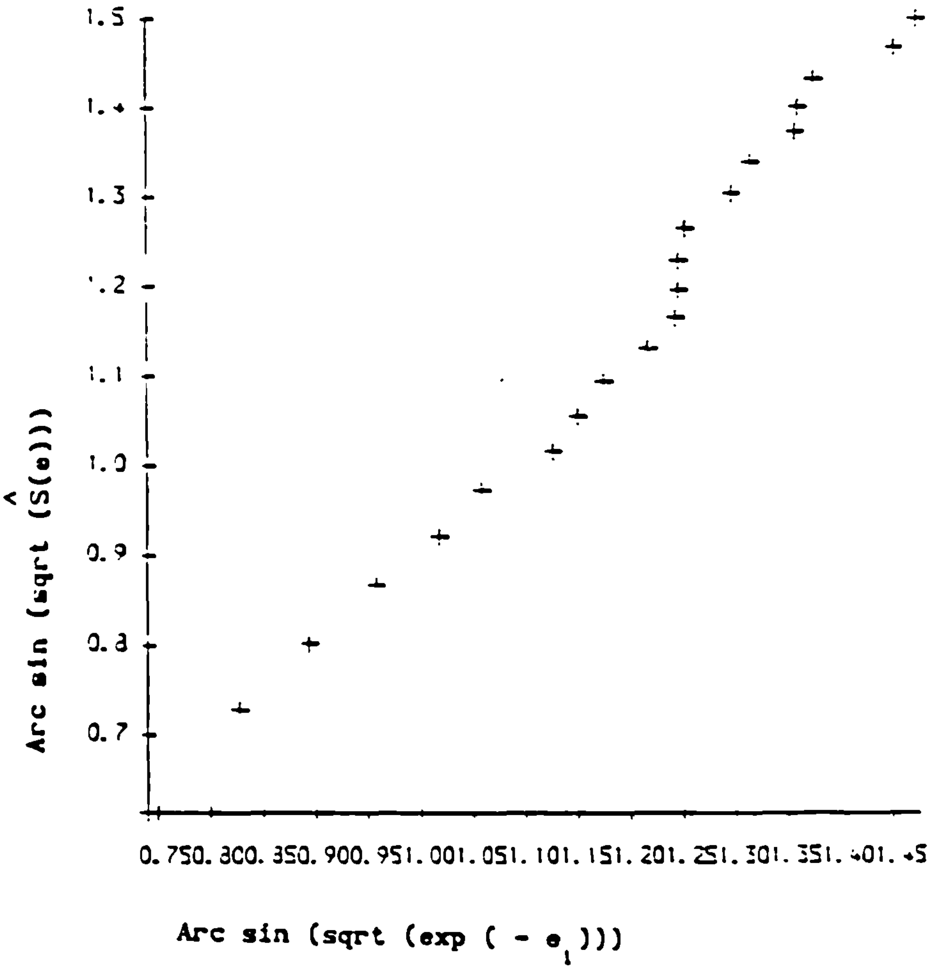


Figure 6.12 Residual plot for log-logistic regression model for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

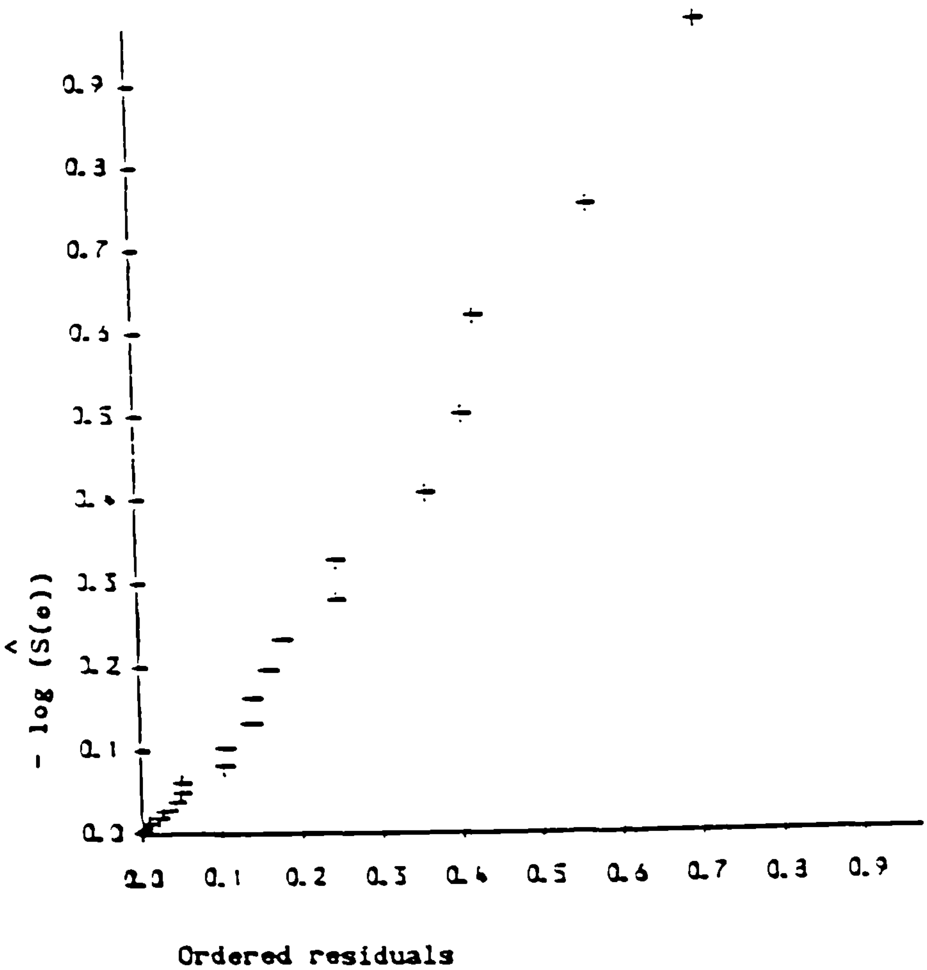


Figure 6.13 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using log-logistic regression model.

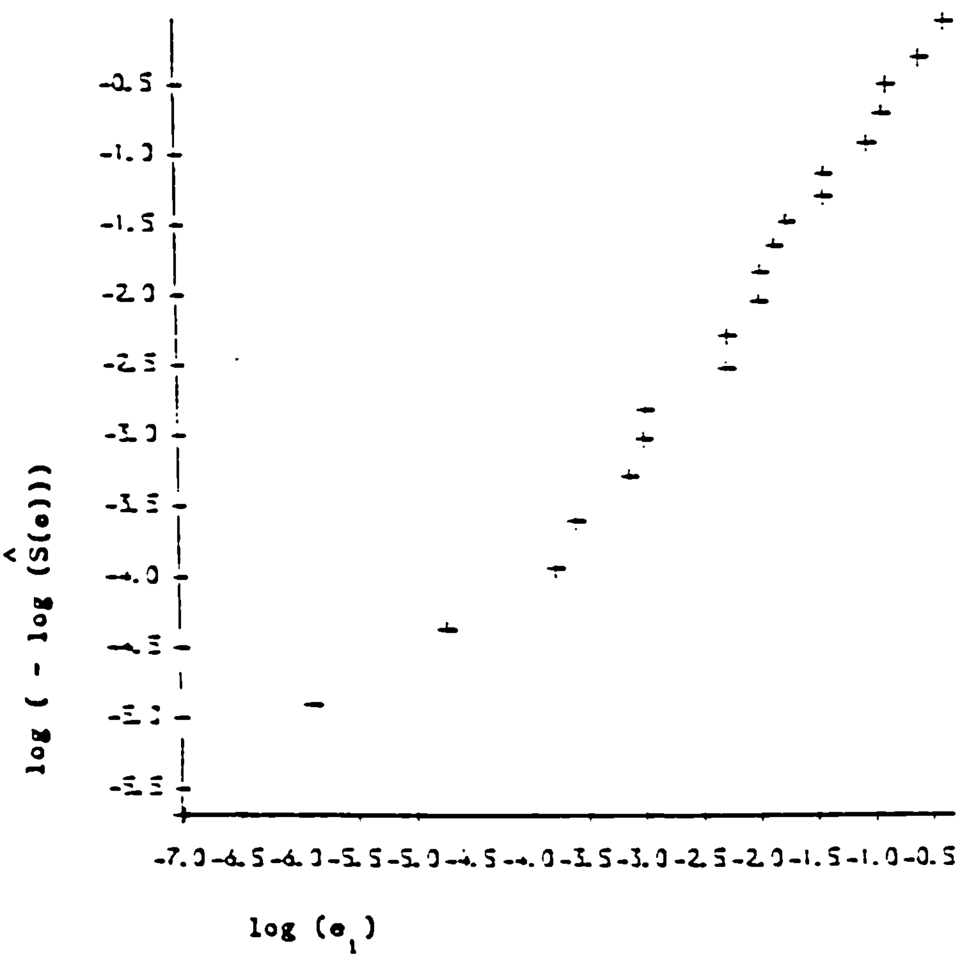
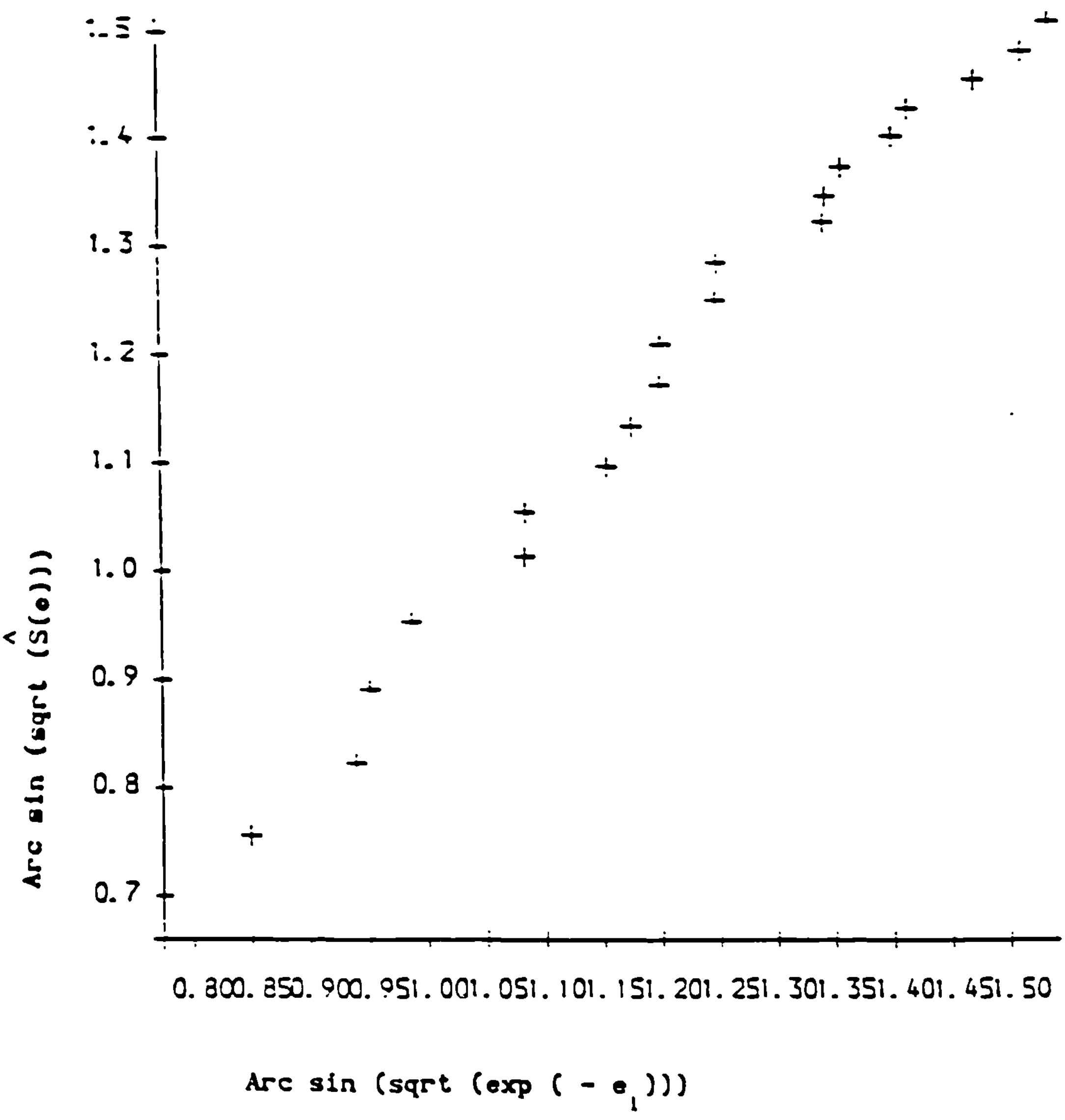


Figure 6.14 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group) using log-logistic regression model.



The results for case a(2) analysis are summarised in Table 6.2.

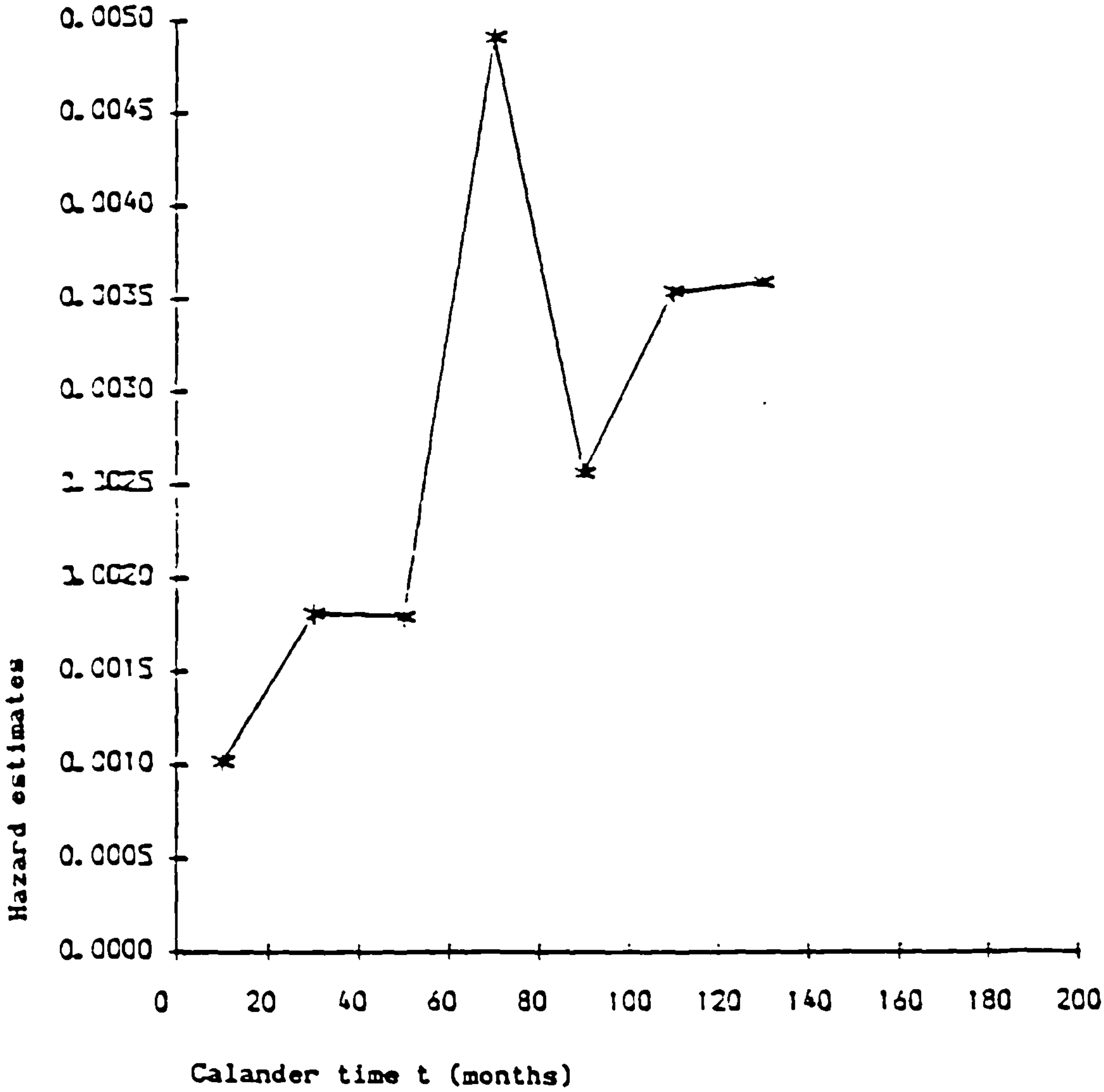
From Table 6.2, it is clear that the Weibull regression model gives a better fit than the other models on account of larger maximised log-likelihood (-256.37). There is as well a significant prognostic effect of the variables FF/C.LIB, CA/C.LIB and C.LIB/TA on survival time with NI/NW and NW/S being of marginal significance. However, this analysis indicates that there is no apparent dependence of survival time of U.K. industrial companies on the variables FF/TA, EBIT/S, QA/TA, FF/S, S/TA, TA/NW, RE/TA and QA/C.LIB.

The survival function estimate $\hat{S}(t)$ in this case, estimated using the Kaplan-Meier method, is plotted in Figure 6.15. $\hat{S}(t)$ indicates a high survival rate but not as high as case a(1). Also the hazard function, useful in modeling survival time data, provides information on how the failure rate will change with the amount of time on test. Figure 6.16 shows that the hazard function has an increasing trend which reaches its peak at approximately 70 months then fluctuates. The peak value is higher than that for case a(1) but it occurs after the same period of time.

Table 6.2 Asymptotic likelihood inference for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using Weibull, exponential and log-logistic regression models .

Model	**Loglike-lihood**	**vari-ables**	**D.F.**	**Coeff-icient**	**S.E.**	**Chi-Sq.**	**P-value**
		intercept	1	6.21	0.55	128.42	0.000
		FF/TA	1	-4.65	3.30	1.98	0.159
		NI/NW	1	0.81	0.44	3.38	0.066
		EBIT/S	1	-0.15	1.21	0.02	0.902
		QA/TA	1	1.49	1.11	1.81	0.179
		FF/S	1	-2.71	2.90	0.88	0.349
		NW/S	1	-0.35	0.20	2.98	0.084
Weibull	-256.37	S/TA	1	0.14	0.13	1.19	0.276
		TA/NW	1	-0.04	0.03	1.76	0.185
		FF/C.LIB	1	3.08	1.04	8.70	0.003
		RE/TA	1	-0.52	0.43	1.45	0.228
		CA/C.LIB	1	-0.47	0.24	4.02	0.045
		QA/C.LIB	1	0.31	0.41	0.55	0.457
		C.LIB/TA	1	-1.62	0.79	4.21	0.040
		scale	1	0.63	0.06		
		intercept	1	6.93	0.85	67.31	0.0001
		FF/TA	1	-7.42	5.22	2.02	0.155
		NI/NW	1	1.26	0.67	3.56	0.059
		EBIT/S	1	-0.73	1.94	0.14	0.707
		QA/TA	1	2.13	1.72	1.54	0.215
		FF/S	1	-3.12	4.68	0.44	0.506
		NW/S	1	-0.44	0.34	1.73	0.189
Exponen-	-266.62	S/TA	1	0.21	0.20	1.05	0.305
tial		TA/NW	1	-0.06	0.05	1.42	0.233
		FF/C.LIB	1	4.60	1.56	8.65	0.003
		RE/TA	1	-0.69	0.68	1.04	0.308
		CA/C.LIB	1	-0.74	0.37	4.11	0.043
		QA/C.LIB	1	0.55	0.65	0.73	0.392
		C.LIB/TA	1	-2.17	1.26	2.96	0.085
		scale	0	1	0		
		intercept	1	6.15	0.60	104.36	0.0001
		FF/TA	1	-4.38	3.51	1.56	0.211
		NI/NW	1	0.77	0.47	2.64	0.104
		EBIT/S	1	-0.11	1.49	0.01	0.942
		QA/TA	1	1.21	1.16	1.09	0.297
		FF/S	1	-3.35	3.35	0.10	0.317
		NW/S	1	-0.36	0.25	2.05	0.152
Log-	-258.44	S/TA	1	0.12	0.12	0.98	0.323
logistic		TA/NW	1	-0.04	0.04	1.02	0.313
		FF/C.LIB	1	3.12	1.09	8.18	0.004
		RE/TA	1	-0.41	0.47	0.78	0.377
		CA/C.LIB	1	-0.55	0.26	4.53	0.033
		QA/C.LIB	1	0.42	0.44	0.90	0.342
		C.LIB/TA	1	-1.55	0.85	3.32	0.069
		scale	1	0.58	0.05		

Figure 6.16 The hazard function estimate for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).



The residual plots for Weibull regression model are given in Figures 6.17, 6.18 and 6.19. The residual plot of e_i against $-\log \hat{S}(e)$, shown in Figure 6.17, suggests the existence of three outliers out of 95 points and this would be expected in the basis of chance alone. However, on the variance stabilized transformed plot in Figure 6.19, these three points (nearest the origin) do not look aberrant. The slopes of the lines based on regression analysis, fitted to the three scatter plots are 1.14, 0.95 and 1.01 respectively, which are close to one, and this implies that the three plots are close to straight lines with unit slope. Therefore the Weibull regression model appears well supported.

Figures 6.20, 6.21 and 6.22, show the residual plots for exponential regression model. The plots exhibit curvature and the fitted straight lines have slopes 1.70, 1.32 and 1.30, which are different from one. Thus the exponential regression model does not appear to be satisfactory.

Also we can see from Figures 6.23, 6.24 and 6.25, that the assumption for the log-logistic regression model are reasonably satisfied with slopes for the fitted lines being 1.21, 0.90 and 1.01. However, these results are not as good as those for Weibull regression model and also the latter is preferable on accounts of its larger maximised likelihood.

Figure 6.17 Residual plot for Weibull regression model for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

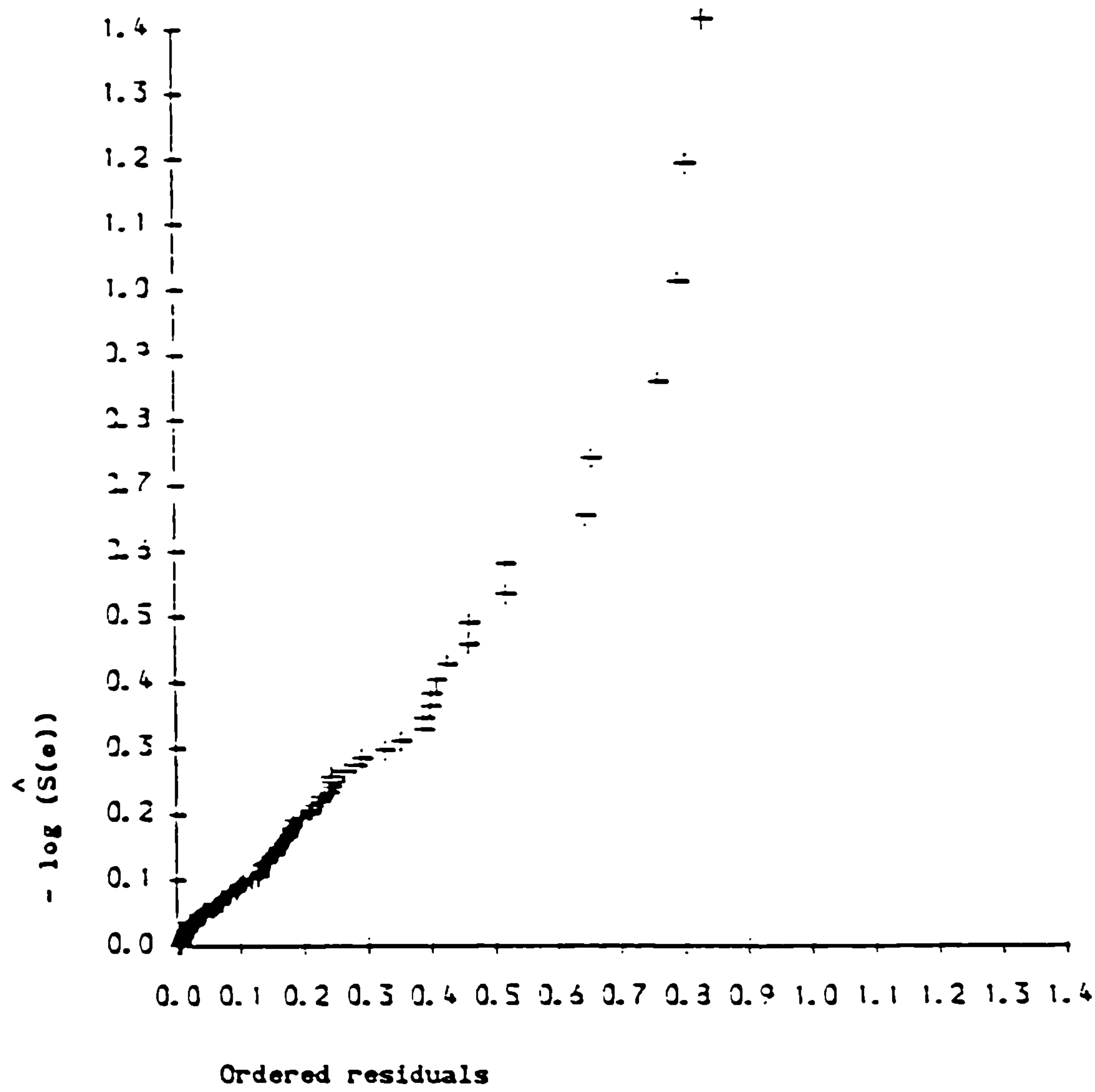


Figure 6.18 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using Weibull regression model.

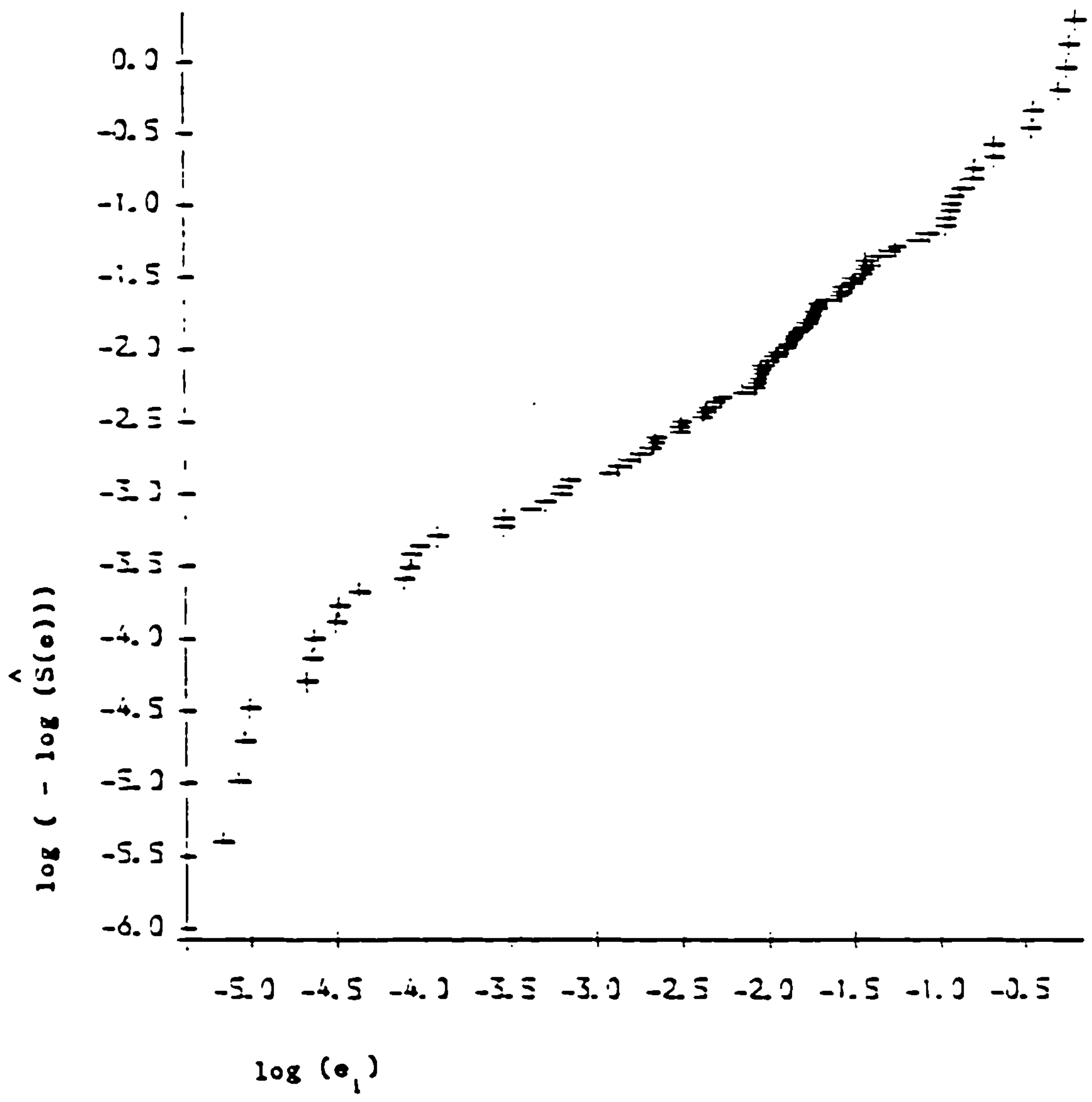


Figure 6.19 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using Weibull regression model.

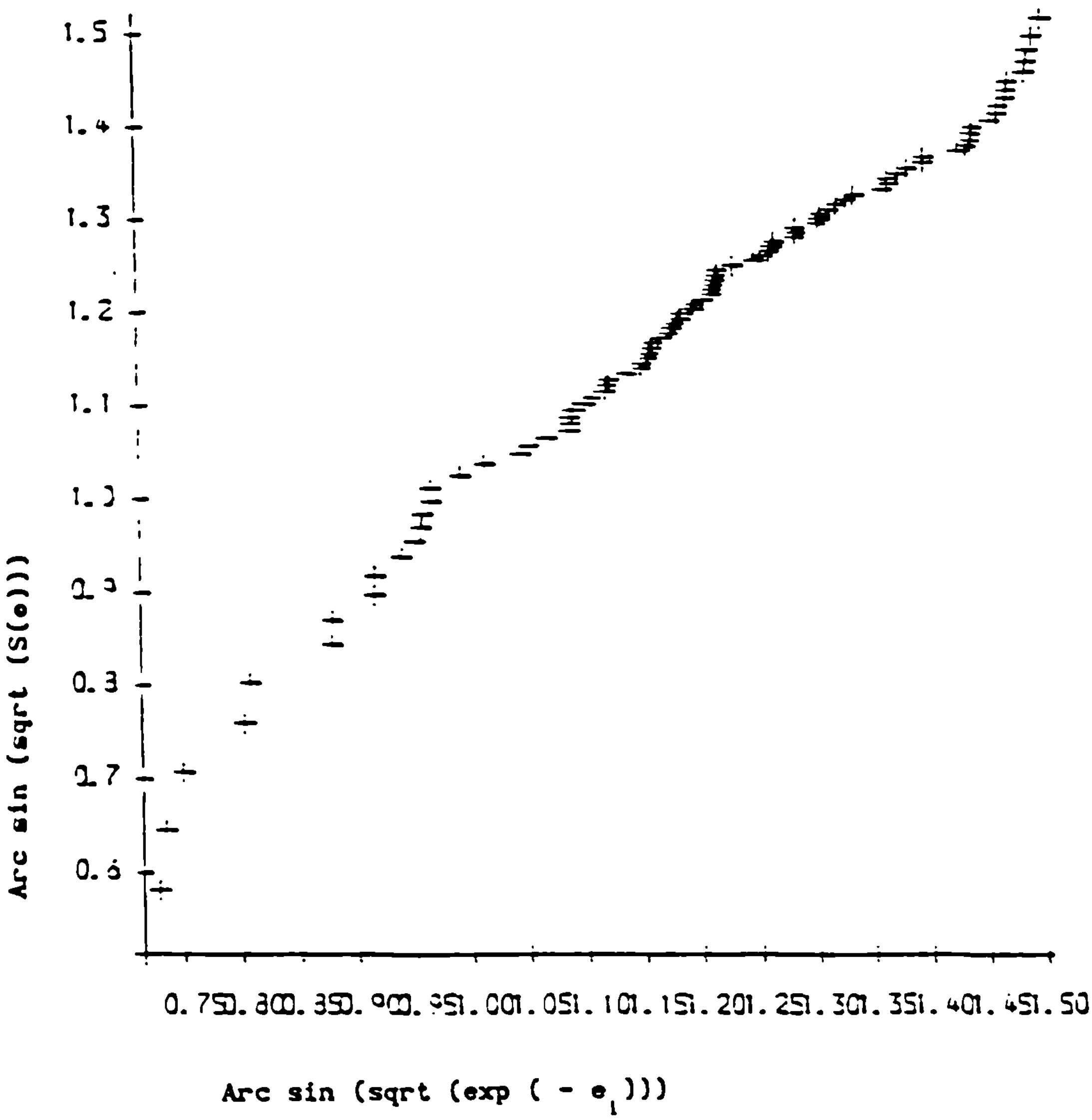


Figure 6.20 Residual plot for exponential regression model for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

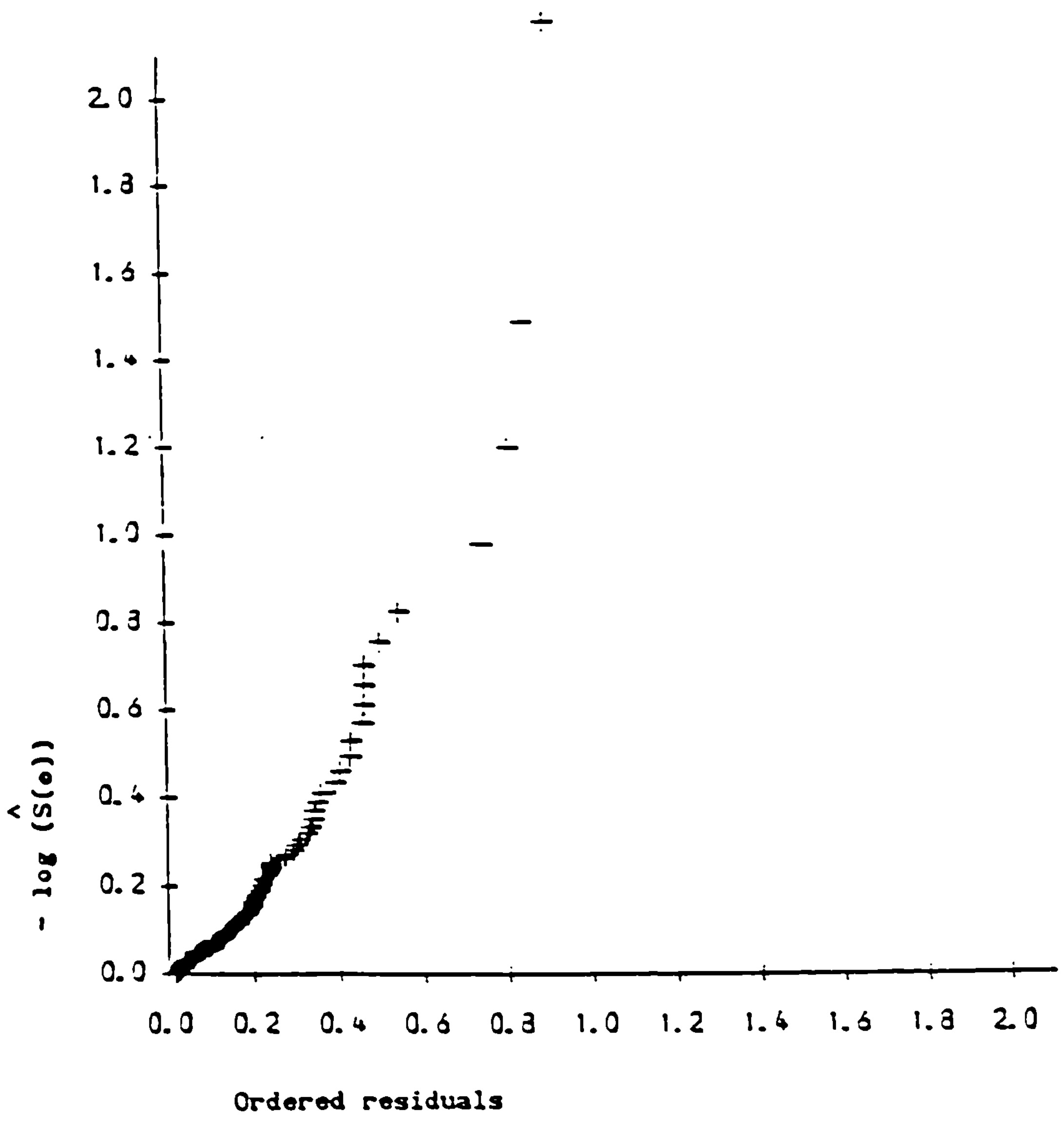


Figure 6.21 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using exponential regression model.

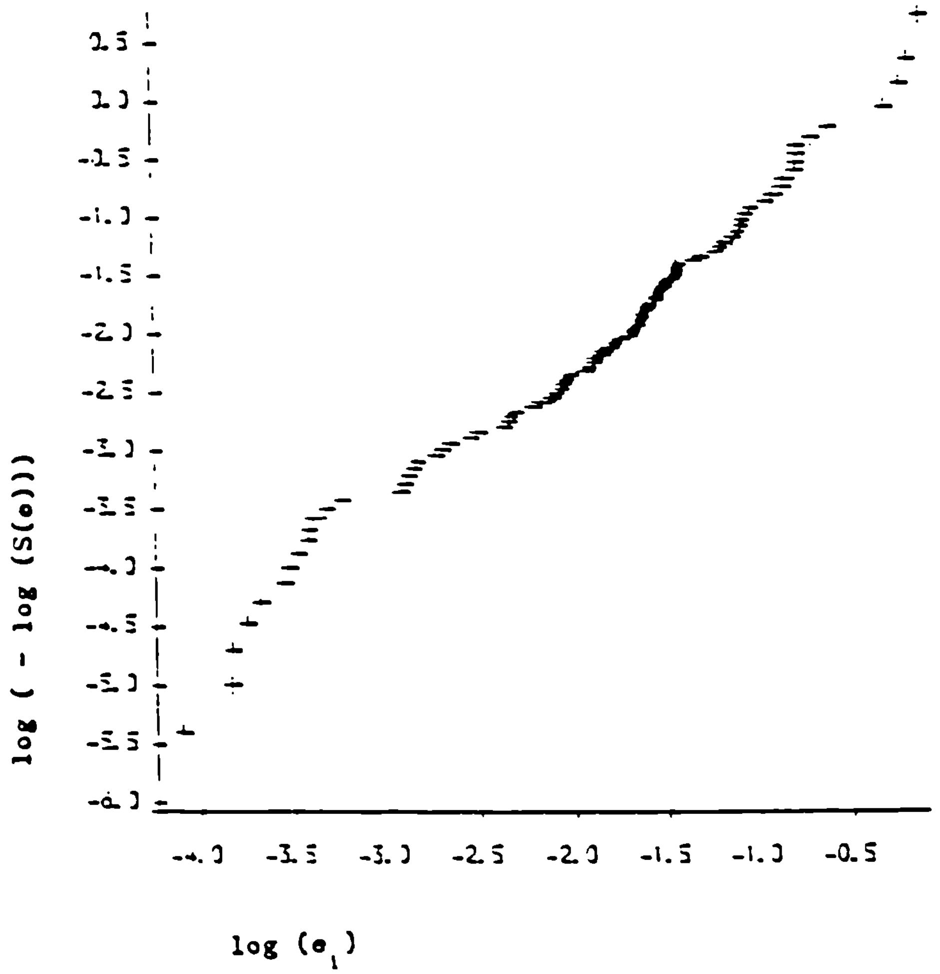


Figure 6.22 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using exponential regression model.

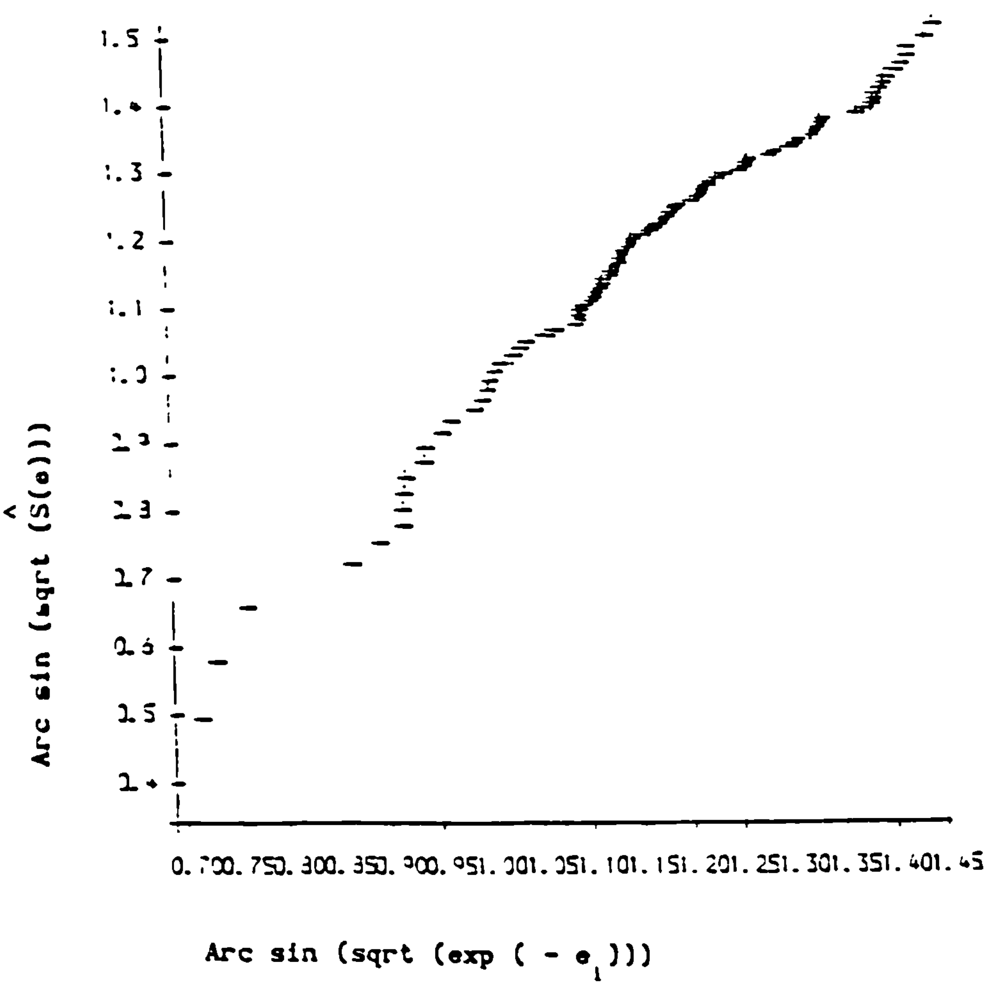


Figure 6.23 Residual plot for log-logistic regression model for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

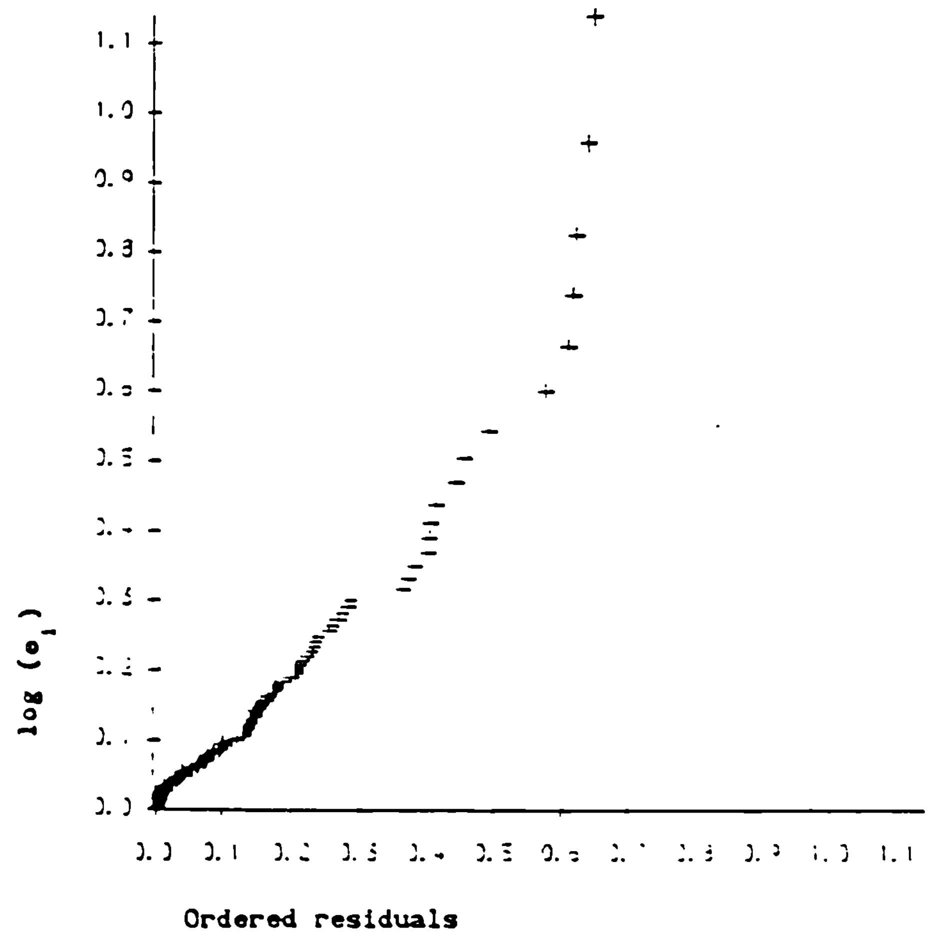


Figure 6.24 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using log-logistic regression model.

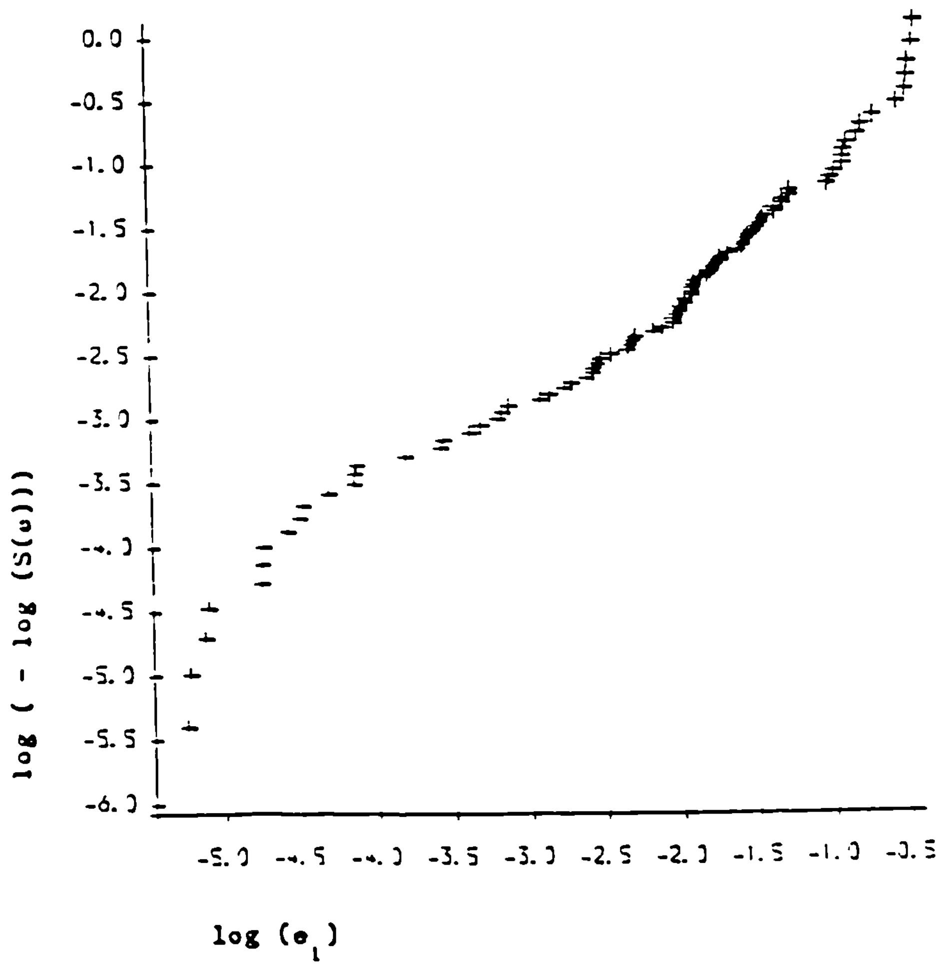
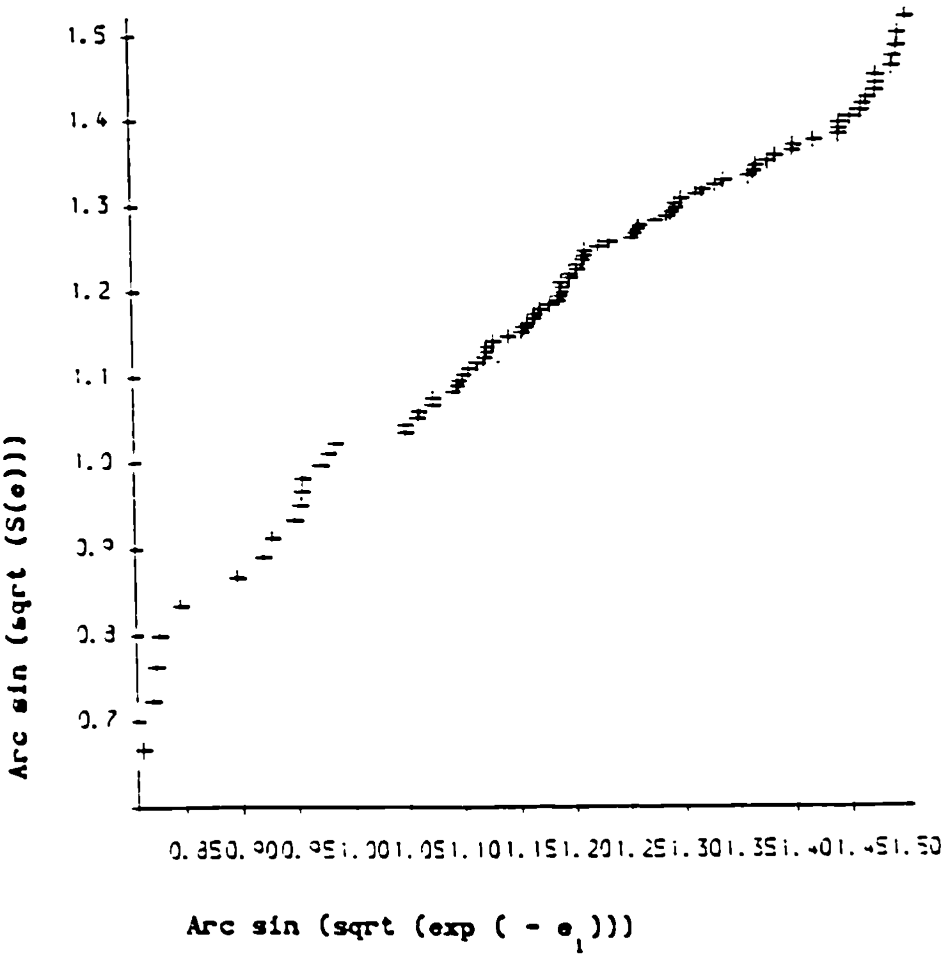


Figure 6.25 Plot of $\sin^{-1}\sqrt{\exp(-e_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case a(2) (the bankrupt , "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group) using log-logistic regression model.



The results of analysis for case b where among the "merged" and "other" group of companies 40 were reclassified as non-survivor and so added to the bankrupt group and the other 34 were reclassified as survivor and hence added to the surviving group are summarised in Table 6.3.

It is clear from the Table that the Weibull regression model gives a better fit than the other models on account of the larger maximised log-likelihood (-131.03). There is as well a significant prognostic effect of the variables CA/TA, TA/NW, FF/C.LIB, RE/TA CA/C.LIB, QA/C.LIB and C.LIB/TA on survival time with FF/TA being of marginal significance. There is no apparent dependence of survival time on the other variables NI/NW, EBIT/S, FF/S, S/TA and CA/S at the 5% significance level. Essentially the same set of variables are significant in all three cases.

The survival function S(t) for the survival time is estimated as before and is plotted in Figure 6.26. It represents a high survival rate, similar to case a(2), but not as high as case a(1). The hazard function in Figure 6.27 shows an increasing trend with some fluctuates.

The residual plots for the Weibull, exponential and log-logistic regression models are represented in Figures 6.28, to 6.36. The plots are very close to straight lines with slopes approximately equal to one (1.13, 0.94, 1.03) for the Weibull regression model so its assumption seems to be well supported and so the model provides a good fit to the data. However, the other two models seem to be less satisfactory (slopes = 1.61, 1.53 , 1.39; and 1.22, 0.90, 1.09 for exponential and log-logistic respectively).

Table 6.3 Asymptotic likelihood inference for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using Weibull, exponential and log-logistic regression models.

Model	**Loglike-lihood**	**Vari-ables**	**D.F.**	**Coeff-icient**	**S.E.**	**Chi-Sq.**	**p-value**
		intercept	1	4.99	0.43	136.07	0.0001
		FF/TA	1	-5.65	3.14	3.23	0.073
		NI/NW	1	0.54	0.36	2.24	0.135
		EBIT/S	1	-0.53	0.92	0.34	0.562
		FF/S	1	-2.06	2.35	0.77	0.379
		CA/TA	1	1.67	0.78	4.54	0.033
Weibull	-131.03	S/TA	1	0.35	0.22	2.67	0.102
		TA/NW	1	-0.05	0.02	4.63	0.032
		FF/C.LIB	1	4.12	1.00	16.81	0.0001
		RE/TA	1	-0.78	0.36	4.76	0.029
		CA/C.LIB	1	-0.60	0.26	5.45	0.020
		QA/C.LIB	1	0.85	0.28	9.12	0.003
		C.LIB/TA	1	-1.60	0.61	6.76	0.009
		CA/S	1	-0.33	0.25	1.78	0.183
		scale	1	0.40	0.042		
		intercept	1	5.42	1.00	29.13	0.0001
		FF/TA	1	-14.81	7.35	4.06	0.044
		NI/NW	1	1.41	0.78	3.28	0.070
		EBIT/S	1	-2.44	2.20	1.22	0.269
		FF/S	1	-0.71	5.56	0.02	0.898
		CA/TA	1	3.47	1.81	3.65	0.056
Expone-ntial	-158.45	S/TA	1	0.98	0.54	3.32	0.069
		TA/NW	1	-0.10	0.05	4.30	0.038
		FF/C.LIB	1	8.88	2.18	16.58	0.0001
		RE/TA	1	-1.55	0.85	3.35	0.067
		CA/C.LIB	1	-1.46	0.59	6.24	0.013
		QA/C.LIB	1	2.06	0.65	10.18	0.001
		C.LIB/TA	1	-3.43	1.42	5.84	0.016
		CA/S	1	-0.43	0.62	0.48	0.490
		scale	0	1	0		
		intercept	1	4.94	0.44	125.34	0.0001
		FF/TA	1	-5.92	3.01	3.87	0.049
		NI/NW	1	0.52	0.33	2.50	0.114
		EBIT/S	1	-0.78	0.10	0.62	0.433
		FF/S	1	-1.67	2.50	0.45	0.505
		CA/TA	1	1.73	0.79	4.74	0.030
Log-logistic	- 133.13	S/TA	1	0.43	0.22	3.78	0.052
		TA/NW	1	-0.04	0.02	3.15	0.076
		FF/C.LIB	1	4.21	1.01	17.29	0.0001
		RE/TA	1	-0.77	0.37	4.34	0.037
		CA/C.LIB	1	-0.72	0.27	6.94	0.008
		QA/C.LIB	1	0.95	0.30	9.96	0.002
		C.LIB/TA	1	-1.96	0.68	8.34	0.004
		CA/S	1	-0.28	0.26	1.24	0.265
		scale	1	0.35	0.04		

Figure 6.26 Kaplan-Meier survival distribution function estimate for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group).

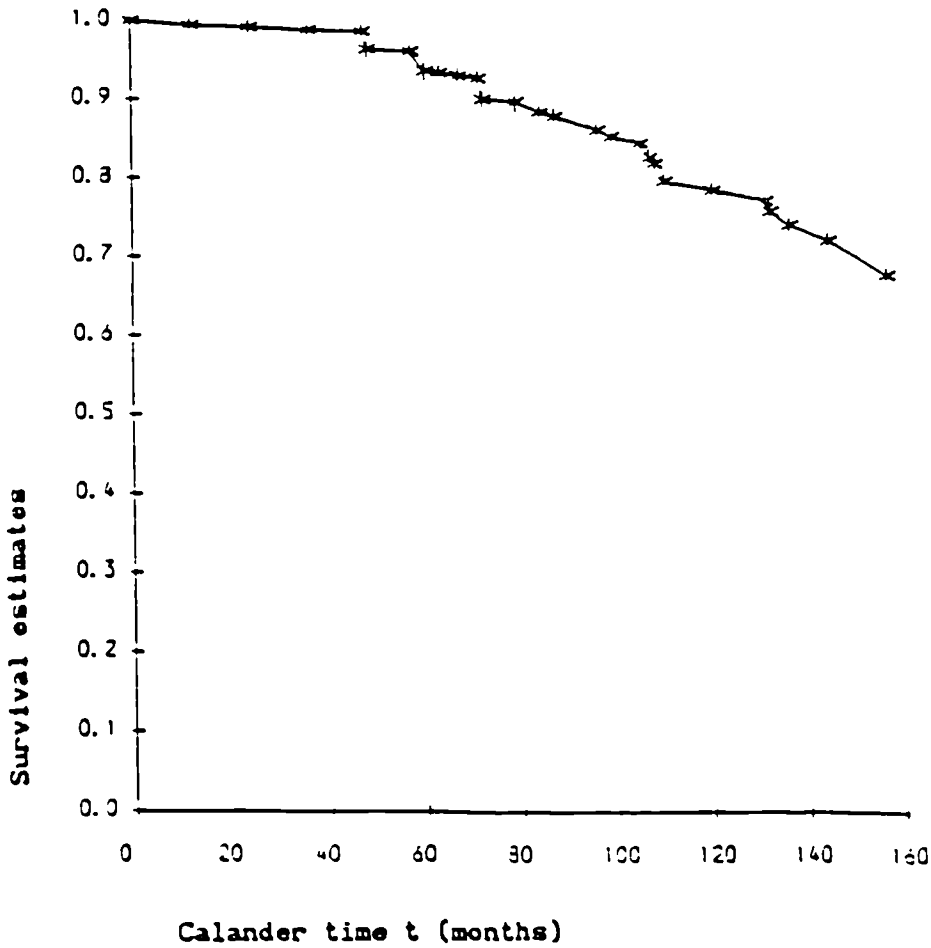


Figure 6.27 The hazard function estimate for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group).

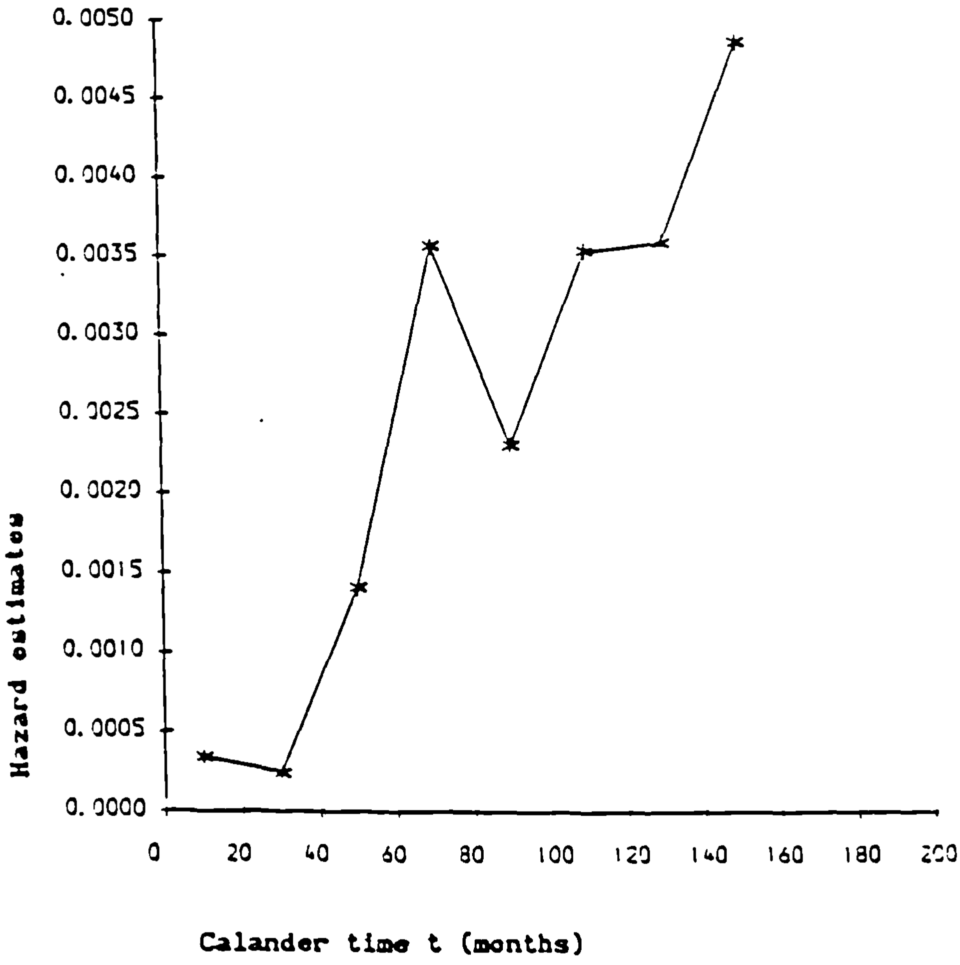


Figure 6.28 Residual plot for Weibull regression model for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group).

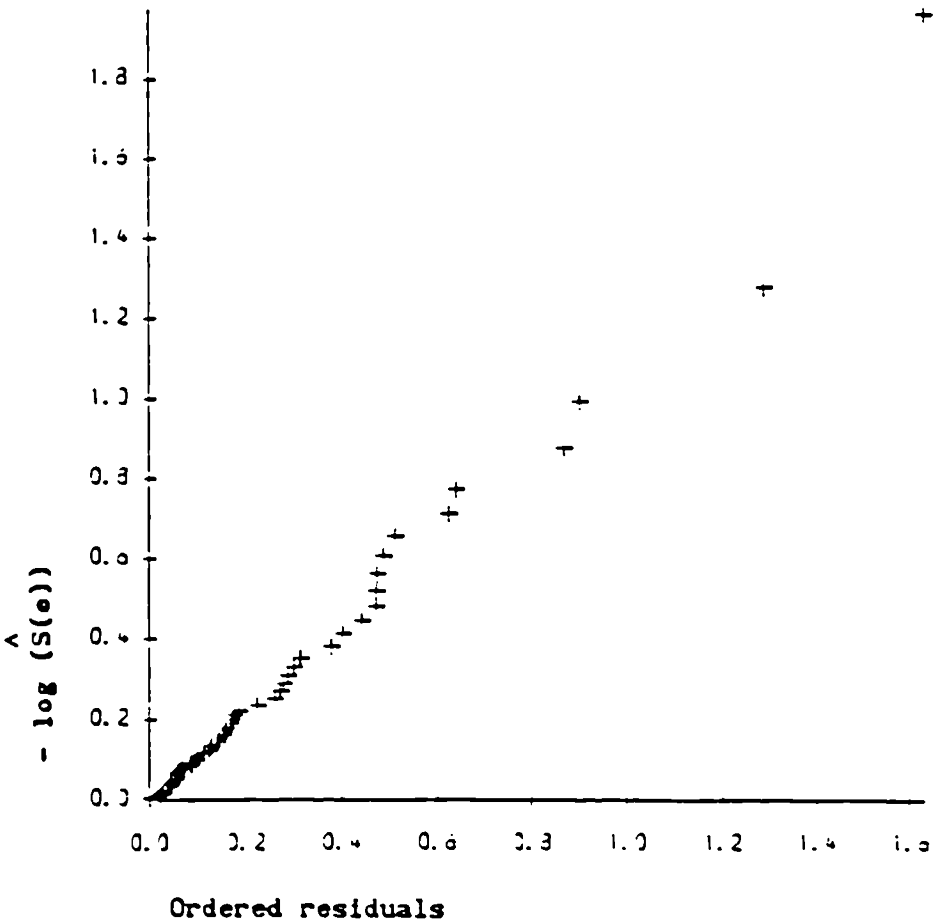


Figure 6.29 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using Weibull regression model.

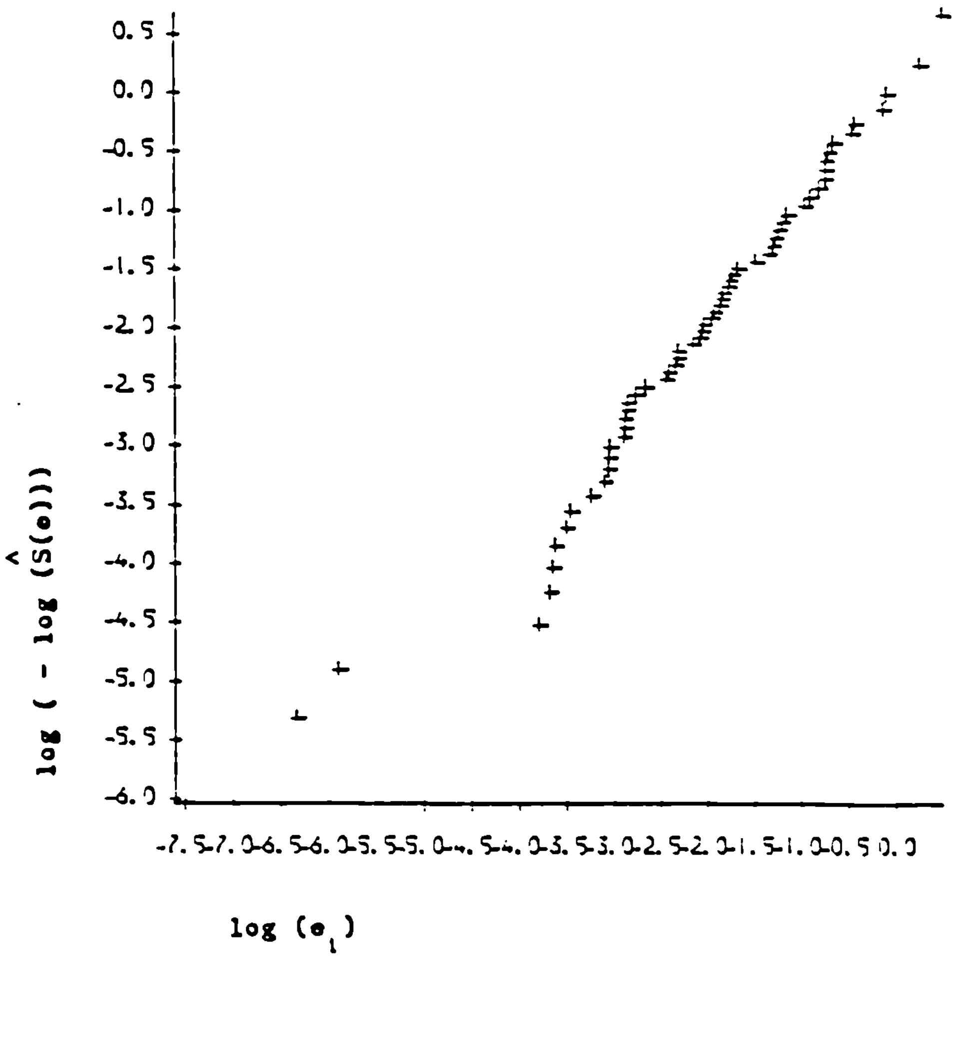


Figure 6.30 Plot of $\sin^{-1}\sqrt{\exp(-\hat{e}_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using Weibull regression model.

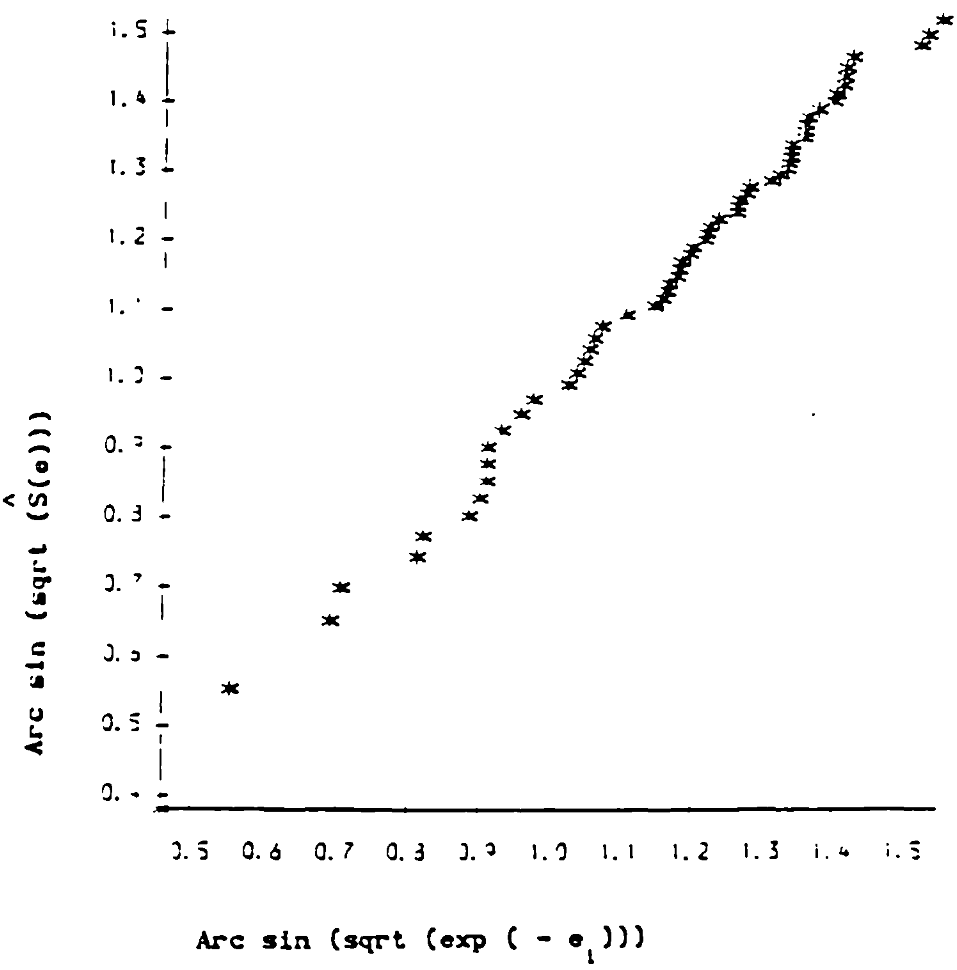


Figure 6.31 Residual plot for exponential regression model for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group).

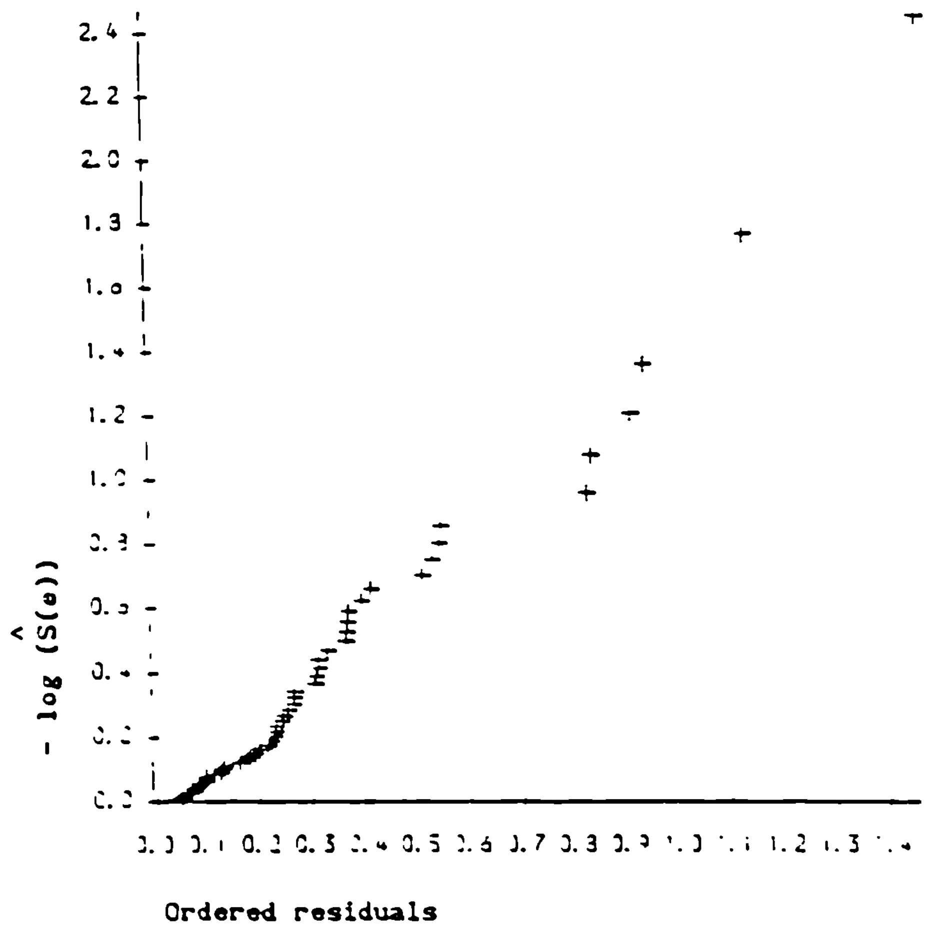


Figure 6.32 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using exponential regression model.

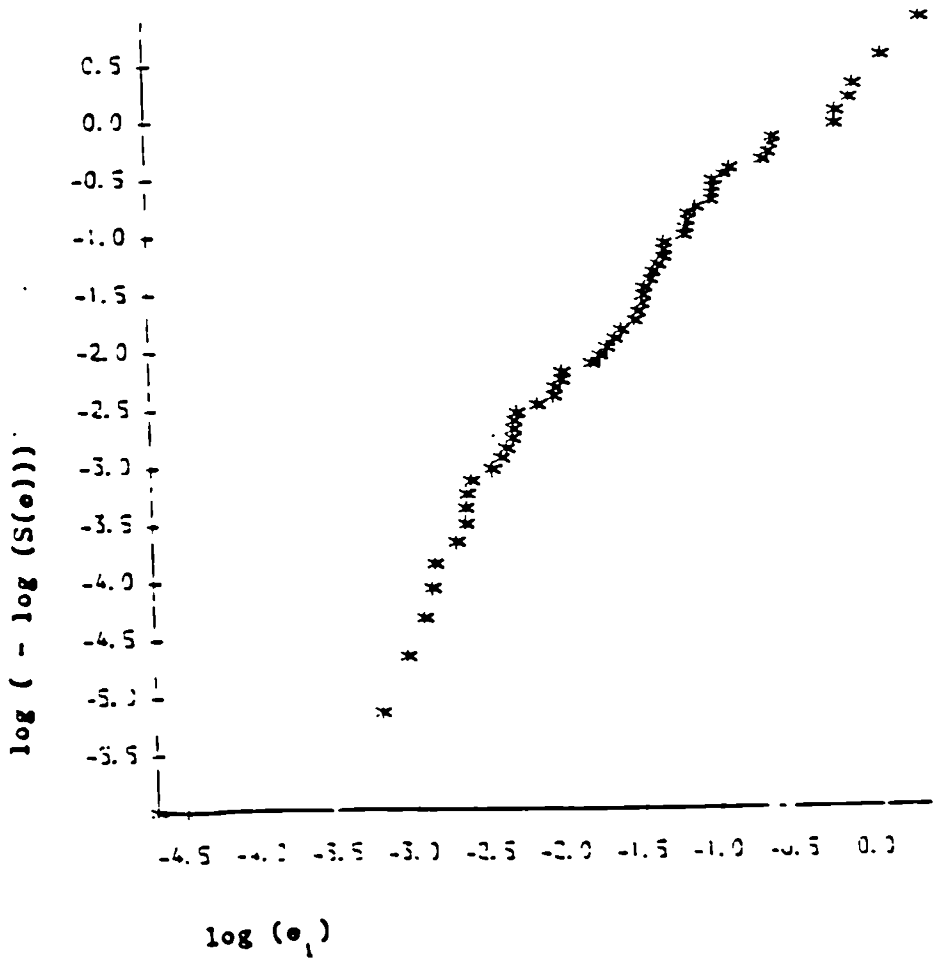


Figure 6.33 Plot of $\sin^{-1}\sqrt{\exp(-\hat{e}_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using exponential regression model.

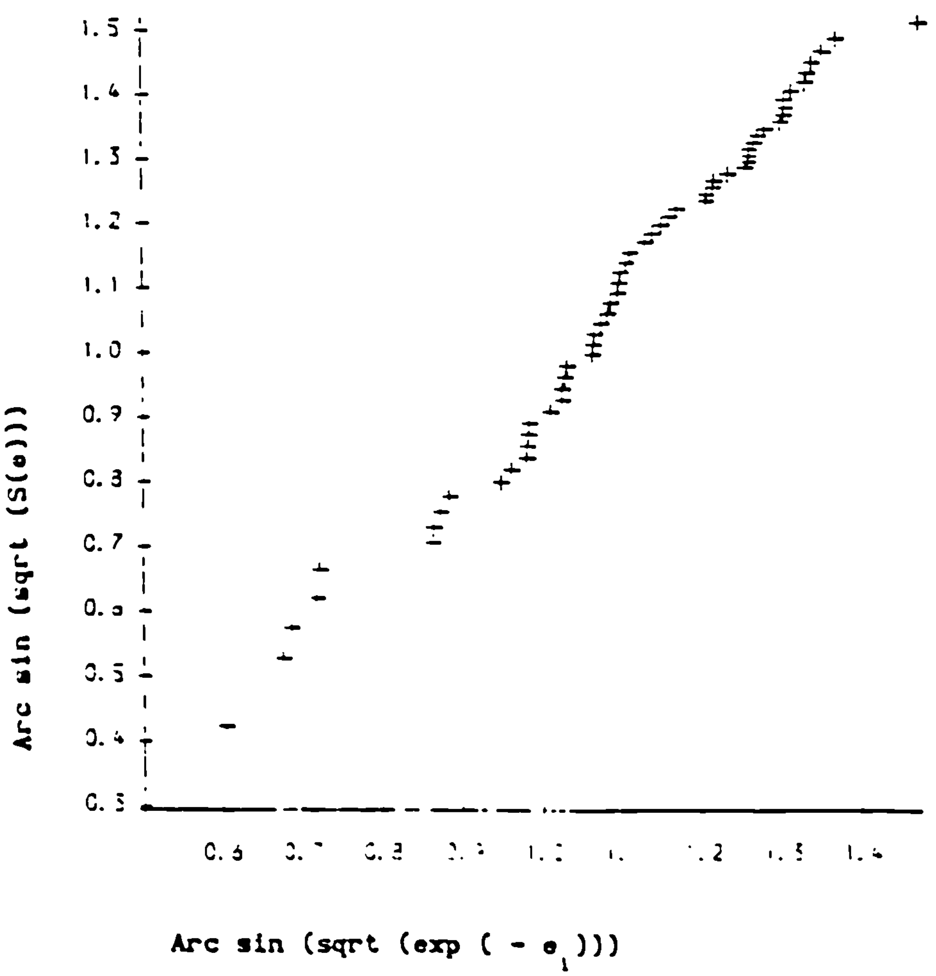


Figure 6.34 Residual plot for log-logistic regression model for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group).

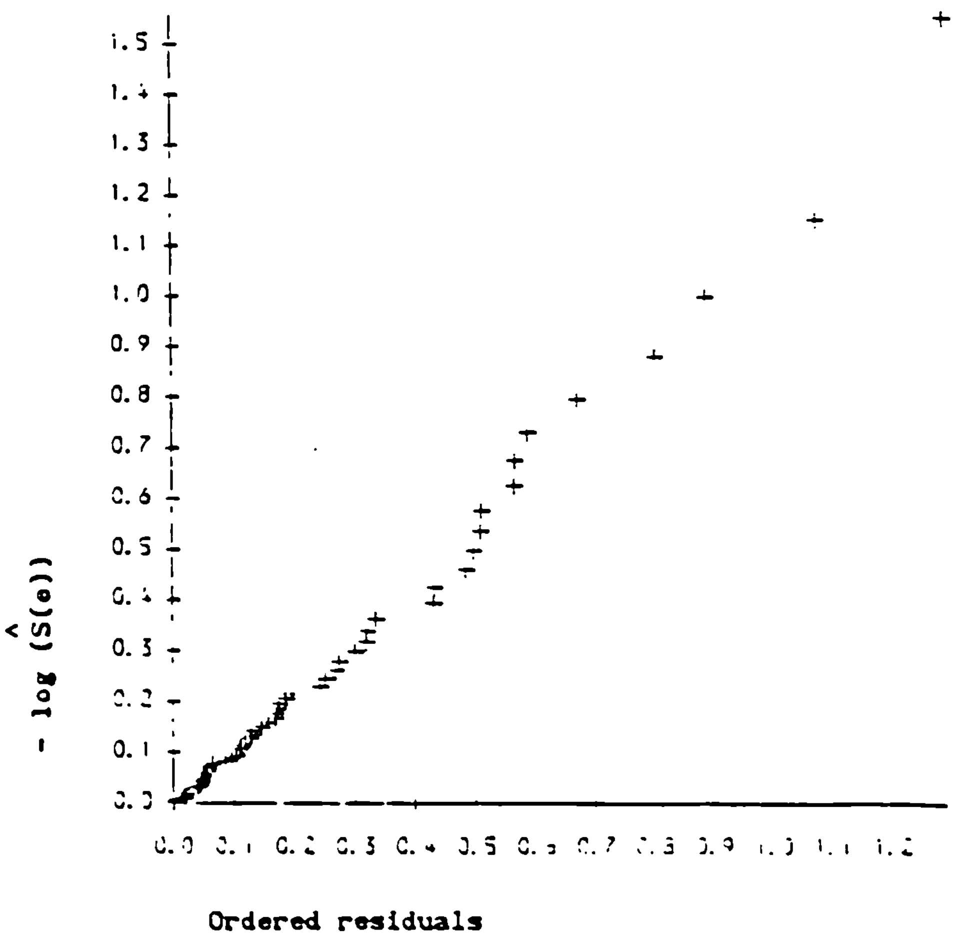


Figure 6.35 Plot of $\log(e_i)$ against $\log(-\log(\hat{S}(e)))$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using log-logistic regression model.

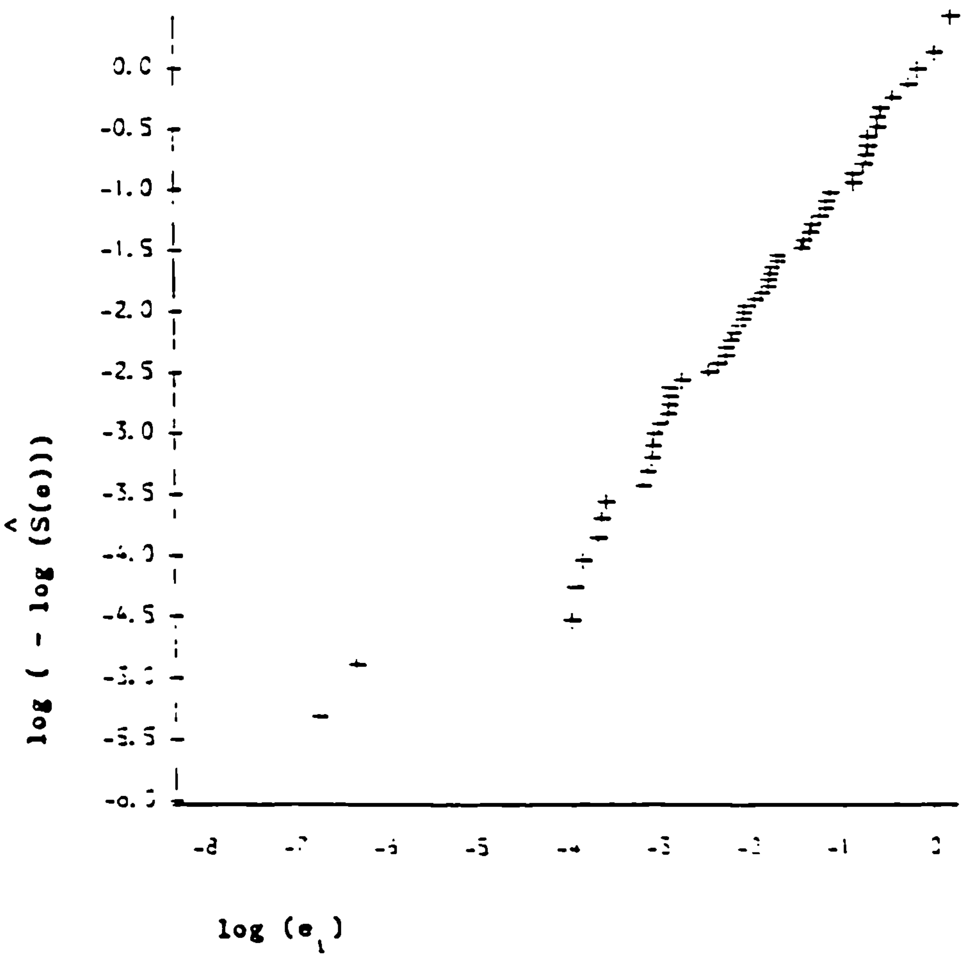
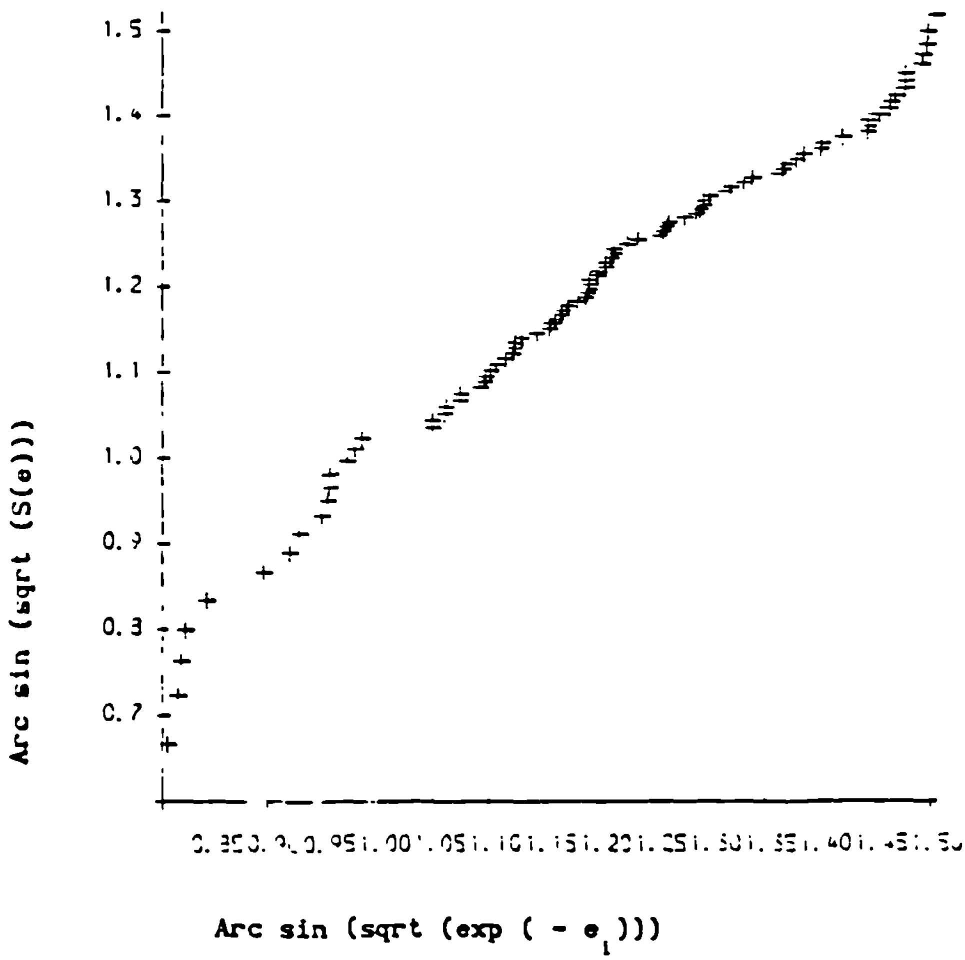


Figure 6.36 Plot of $\sin^{-1}\sqrt{\exp(-\hat{e}_i)}$ against $\sin^{-1}\sqrt{\hat{S}(e)}$ for case b (bankrupt and 40 others companies as a non-surviving group and the surviving and 34 others companies as surviving group) using log-logistic regression model.



6.6.2 Nonparametric Models

The proportional hazard model which has an underlying distribution function was fitted to the data. The results for case a(1) analysis are summarised in Table 6.4. The Table also shows the order of importance for the explanatory variables selected by the stepwise regression procedure of Section 6.5.3.

Table 6.4 Summary results for a proportional hazard model for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

Variable order	**Cofficient**	**S.E.**	**Chi-sq.**	**P - value**
NI/NW	-2.44	0.68	12.99	0.0003
CA/S	1.21	0.58	4.37	0.037
FF/C.LIB	-7.73	2.00	14.89	0.0001
FF/S	15.59	5.36	8.46	0.004

It indicates that NI/NW is the most significant financial ratio in explaining failure, followed by CA/S, FF/C.LIB and FF/S. Eight other explanatory variables FF/TA, EBIT/S, CA/TA, NW/S, RE/TA, CA/C.LIB, QA/C.LIB and C.LIB/TA were not selected by the stepwise procedure and so have been omitted from the Table.

The regression coefficients indicate the relationship between the covariates (explanatory variables) and the hazard function. A positive coefficient increases the value of the hazard function and therefore indicates a negative relationship with survival time and this is true in the case of CA/S and FF/S ratios. A negative coefficient has the reverse interpretation where the value of hazard function is decreased so increasing the value of survival time, which is true in the case of the ratios NI/NW and FF/C.LIB.

The residual plot of the proportional hazard model is represented in Figure 6.37. This plot is approximately straight line with slope = 1.03.

The hazard function in Figure 6.38, shows the hazard estimated from the model for the four significant variables NI/NW, CA/S, FF/C.LIB and FF/S at 0.5. It shows a mild fluctuating trend which reaches its peak at 5 years and with some evidence that the hazard increases again after about 8 years.

Figure 6.37 Residual plot for proportional hazard model for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

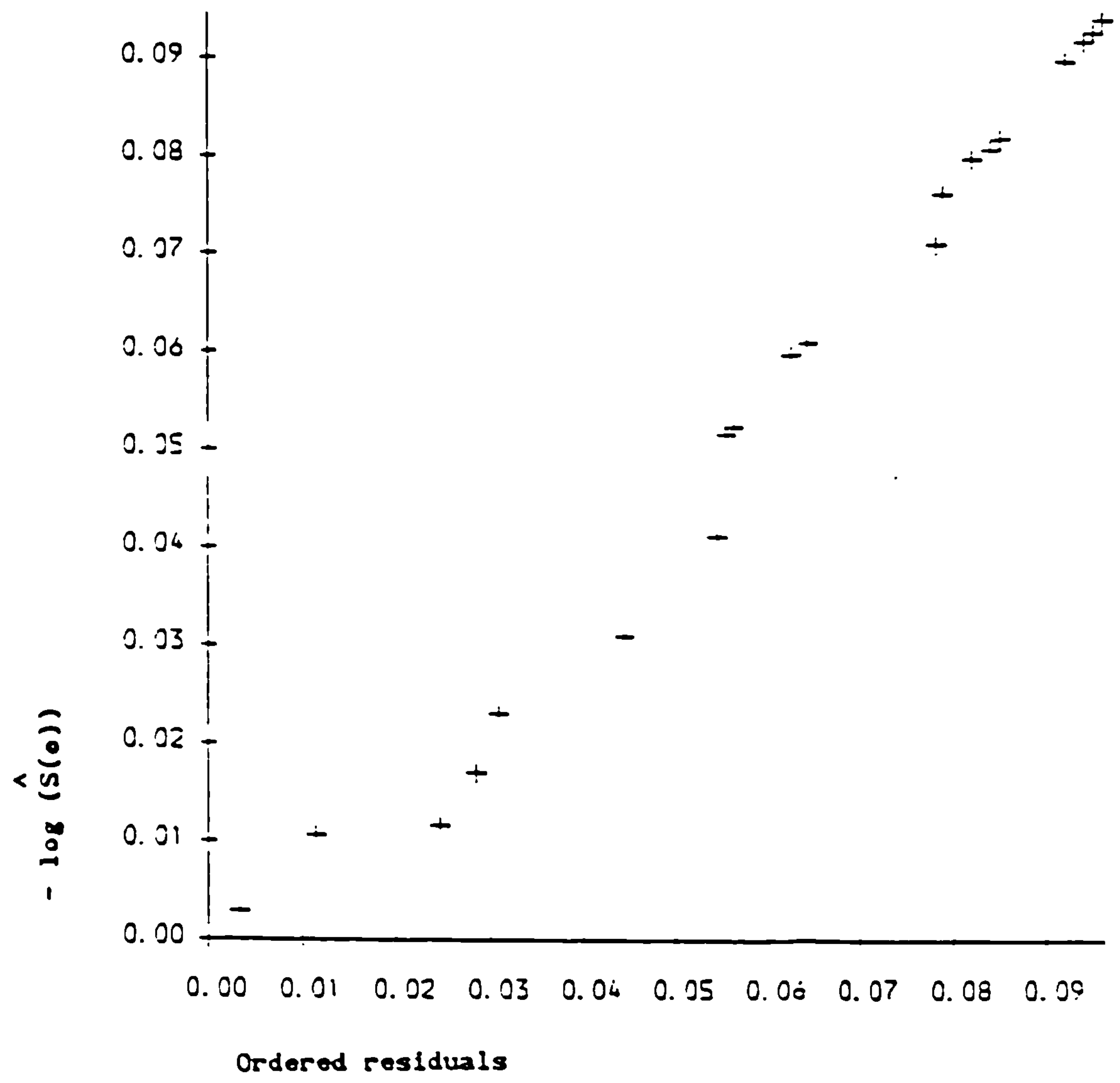


Figure 6.38 The hazard function estimate for case a(1) (the bankrupt companies, treated as a non-surviving group and surviving companies as a second group).

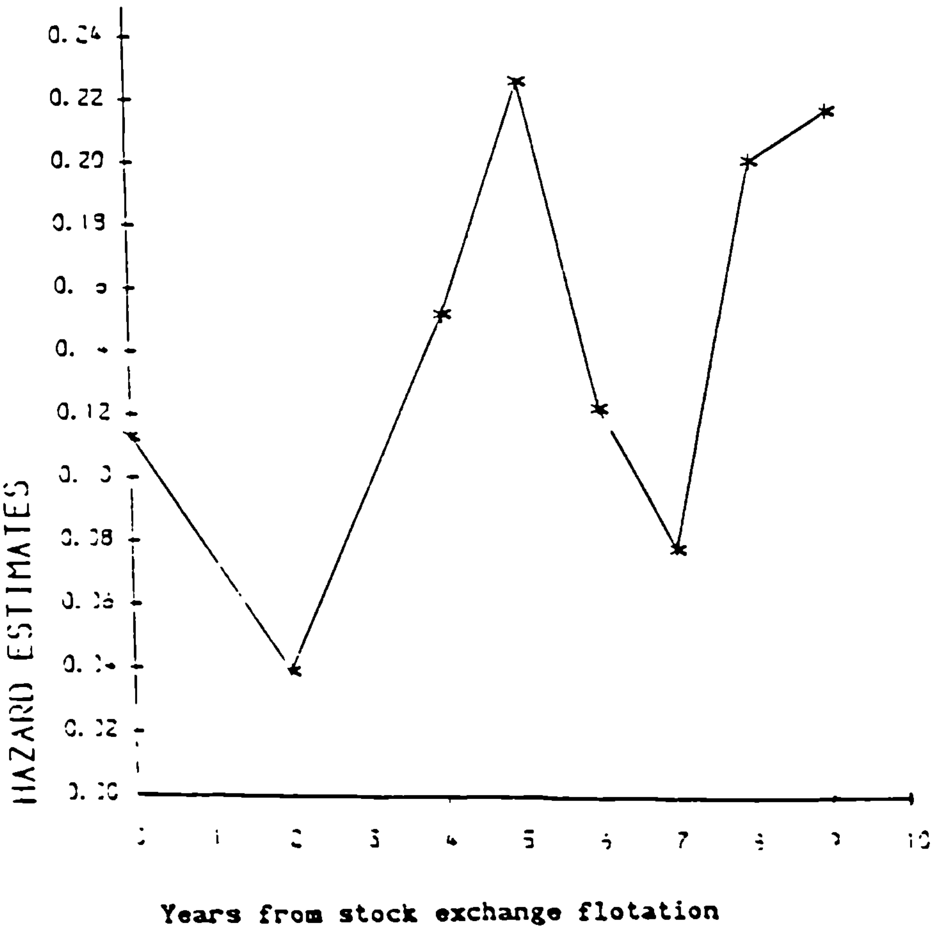


Table 6.5 summarises the results from the proportional hazard model for case a(2) where the "merged" and "other" companies are grouped with the bankrupt ones.

Table 6.5 Summary results for a proportional hazard model for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

Variable order	Coefficient	S.E.	Chi-sq.	P -value
FF/C.LIB	-1.57	0.43	13.45	0.0002
NW/S	0.50	0.16	9.19	0.002

The results from Table 6.5 shows the order of fit where the variable FF/C.LIB is the most significant ratio in explaining failure, followed by NW/S. Eleven other explanatory variables FF/TA, NI/NW, EBIT/S, QA/TA, FF/S, S/TA, TA/NW, RE/TA, CA/C.LIB, QA/C.LIB and C.LIB/TA were not selected by the stepwise procedure and so have been omitted from the Table. Increases in the value of the variable FF/C.LIB decreases the value of hazard function and so increases the value of survival time, while survival time decreases as NW/S increases.

The residual plot in this case is represented in Figure 6.39.

This plot is approximately a straight line and thus the assumption of the model seem well supported, indicating the adequacy of the fit.

The hazard function given in Figure 6.40 shows the hazard estimated from the model which includes FF/C.LIB and NW/S only at 0.5. The Figure shows there is an increase in risk between 3 and 5 years, peaking at 5 years but with some evidence that the hazard increases again after about 8 years.

Figure 6.39 Residual plot for proportional hazard model for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

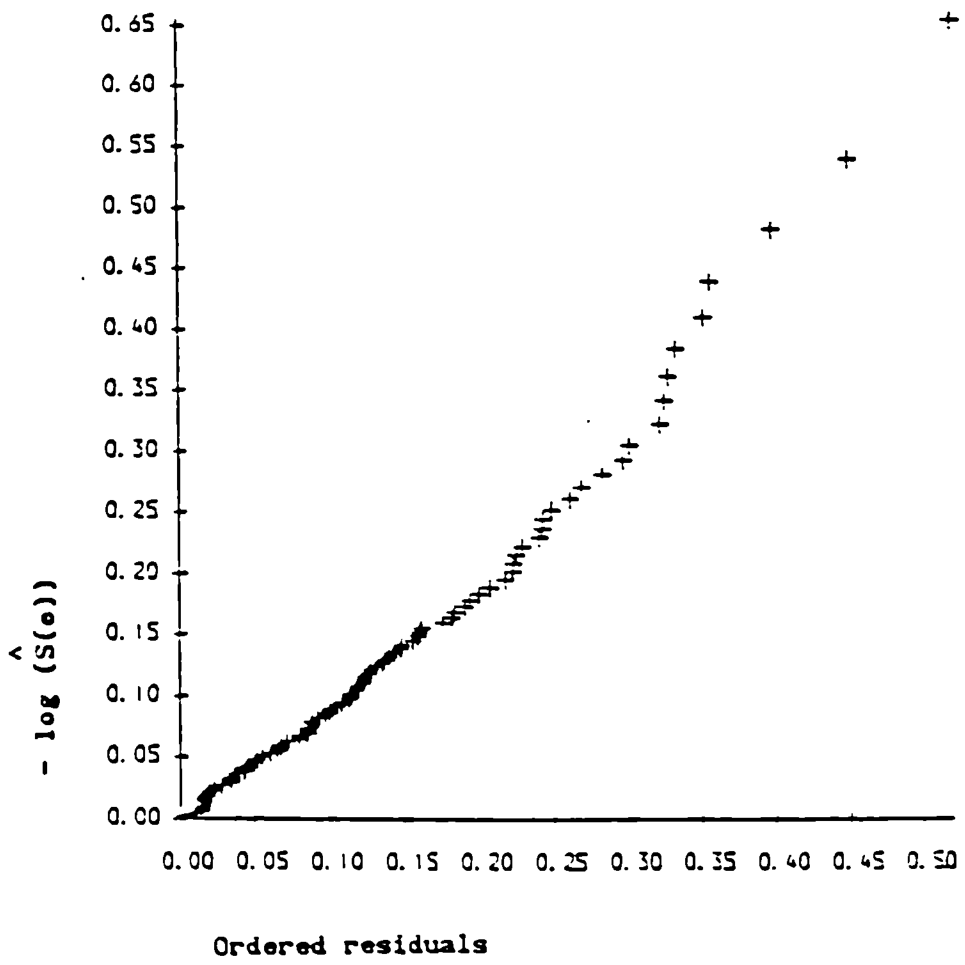


Figure 6.40 The hazard function estimate for case a(2) (the bankrupt, "merged" and "other" companies, treated as a non-surviving group and surviving companies as a second group).

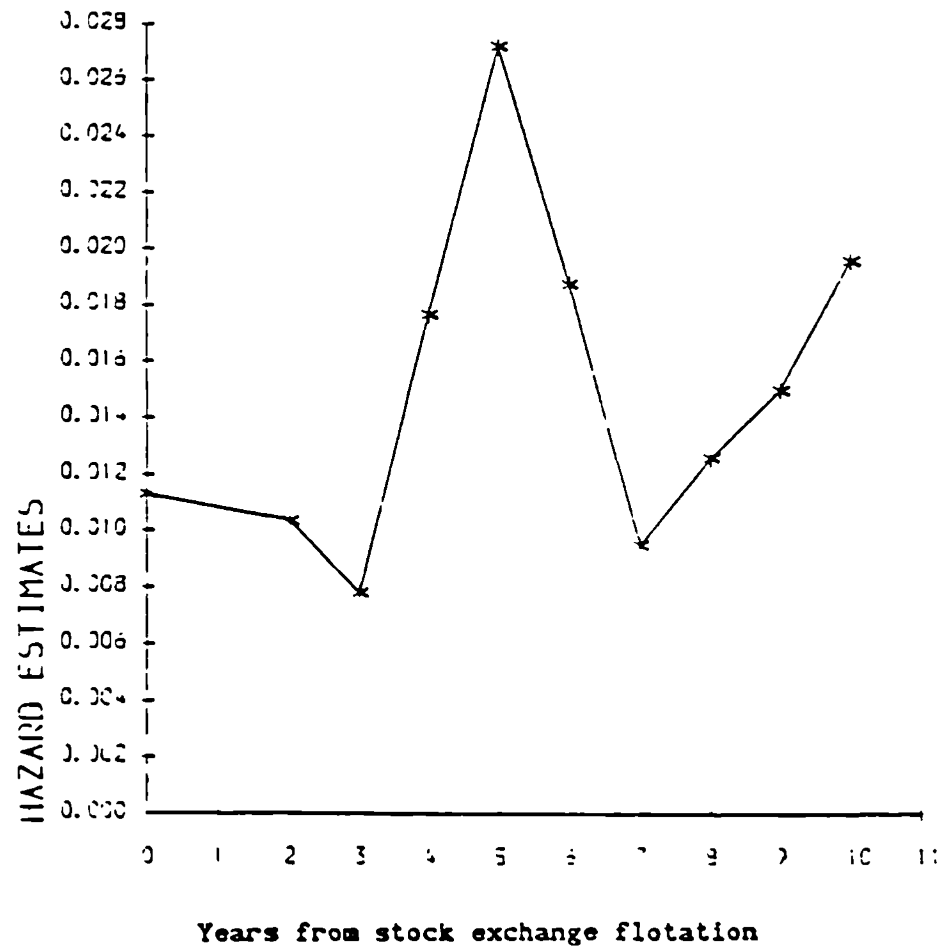


Table 6.6 summarises the results from the proportional hazard model for case b by using a similar modelling procedure as before for the companies *after* reclassifying the "merged" and "other" companies into bankrupt and survivor.

Table 6.6 Summary results for a proportional hazard model for case b (bankrupt and 40 others as a non-surviving group of companies and the surviving and 34 others as a surviving group).

Variable order	**Coefficient**	**S.E.**	**Chi-sq.**	**P-value**
FF/C.LIB	-9.82	2.17	20.45	0.0000
QA/C.LIB	-1.17	0.38	9.30	0.0023
FF/TA	16.64	5.96	7.80	0.0052
S/TA	-1.48	0.37	16.12	0.0001
EBIT/S	4.23	1.29	10.85	0.0010
NI/NW	-1.51	0.71	4.54	0.0331

It indicates that FF/C.LIB is the most significant ratio in explaining failure, followed by QA/C.LIB, FF/TA, S/TA, EBIT/S and NI/NW. Seven other explanatory variables FF/S, CA/TA, TA/NW, RE/TA, CA/C.LIB, C.LIB/TA and CA/S were not selected by the stepwise procedure and so have been omitted from the Table.

Increases in the value of the variables FF/C.LIB, QA/C.LIB, S/TA and NI/NW decreases the value of hazard function and so increases the value of survival time while survival time decreases as FF/TA and EBIT/S increases.

The residual plot in this case is represented in Figure 6.41. This plot is approximately a straight line and thus the assumption of the model seem well supported.

The hazard function given in Figure 6.42 shows the hazard estimated from the model which includes FF/C.LIB, QA/C.LIB, FF/TA, S/TA, EBIT/S and NI/NW at 0.5, 0.5, 0.5, 0.9, 0.5 and 0.5 respectively. The Figure shows there is an increase in risk between 3 and 6 years, peaking at 6 years but with some evidence that the hazard increases again after about 9 years.

Figure 6.41 Residual plot for proportional hazard model for case b (bankrupt and 40 others as a non-surviving group of companies and the surviving and 34 others as a surviving group).

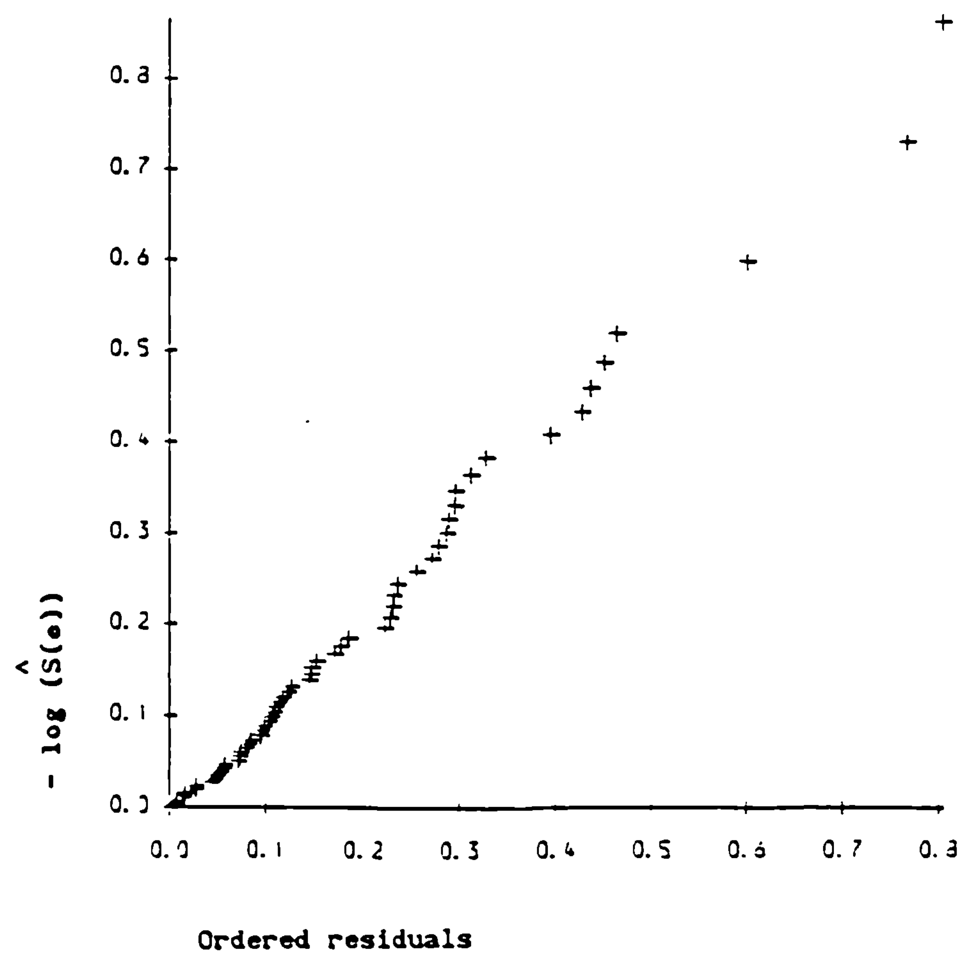
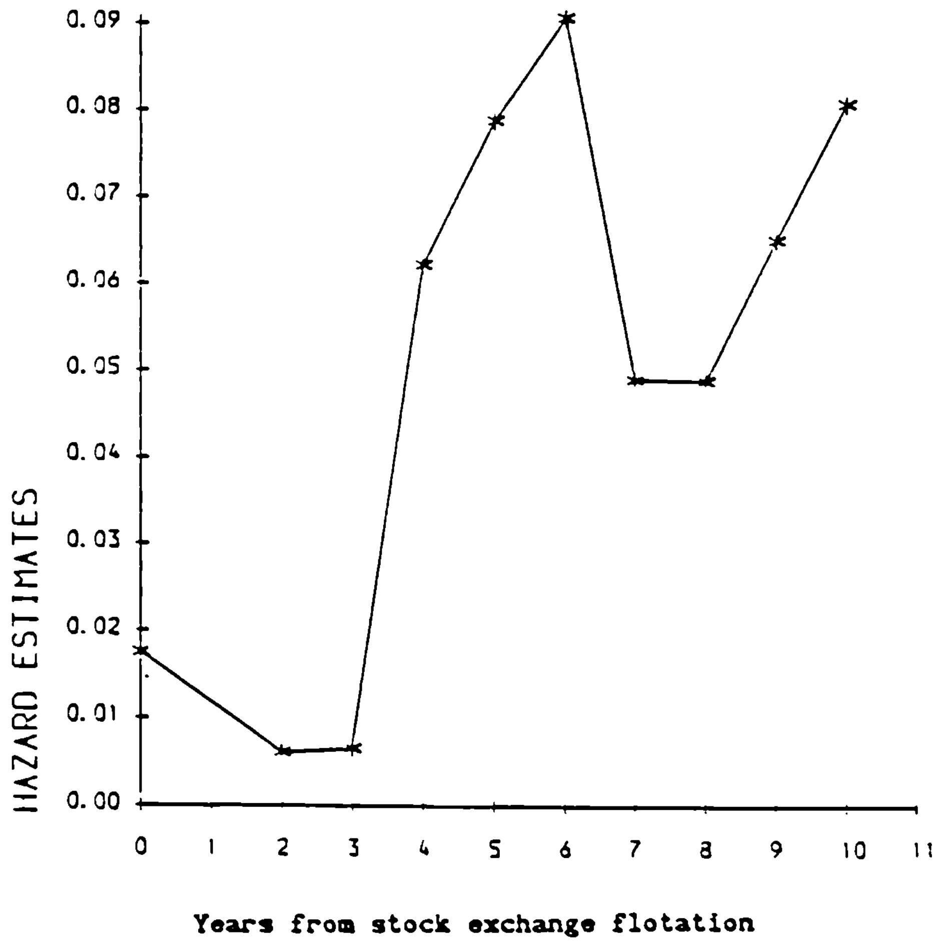


Figure 6.42 The hazard function estimate for case b (bankrupt and 40 others as a non-surviving group of companies and the surviving and 34 others as a surviving group).



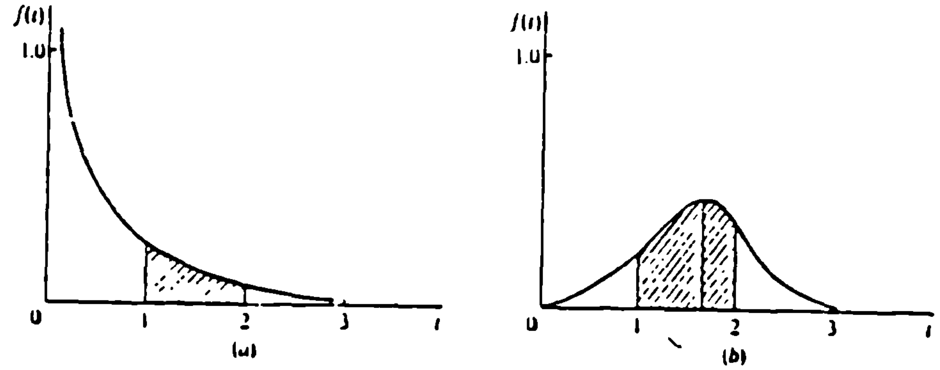
6.6.3 Estimation the Probability of Company Failure

Our second task in this chapter is to estimate failure probabilities for each company over its duration in the study, based on parametric (Weibull regression model) and nonparametric (proportional hazard model) models. The method of estimating the probability of company failure is based on the survival function $S(t)$ and the cumulative distribution function $F(t)$ of the survival time T. The probability density function $f(t)$ of T is defined as the limit of the probability that an individual company fails in the short interval t to $t + \Delta t$ per unit width Δt, or simply the probability of failure in a small interval per unit time. It can be expressed as

$$f(t) = \lim_{\Delta t \to 0} \frac{P\{\text{ an individual dying in the interval } (t, t+\Delta t)\}}{\Delta t}$$

The proportion of companies that fail in any time interval and the peak of frequency of failure can be found from the density function. For example, density curve in Figure 6.43(a) gives a pattern of high failure rate at the beginning of the study which decreases as time increases. While in Figure 6.43(b) the peak occurs at approximately 1.7 units of time. The proportion of companies that fail between 1 and 2 units of time is equal to the shaded area under the density curve.

Figure 6.43 Two examples of the density curve of survival time



Therefore the cumulative distribution function is

$$F(t) = \int_0^t f(u)\, du.$$

The survival function (Section 6.2) is defined as the probability that an individual survives at least time t (t>0) i.e.

$$\begin{aligned} S(t) &= \text{Pr (an individual survives at least time t)} \\ &= \int_t^\infty f(u)\, du = \text{pr}(T \geq t) \end{aligned}$$

Then

$$\begin{aligned} S(t) &= 1 - \text{Pr(an individual fails before time t)} \\ &= 1 - F(t) \end{aligned}$$

Therefore the probability of failure before time t equals $1 - S(t)$.

We estimated $S(t;X)$ for the Weibull regression model as

$$\hat{S}(t;X) = \exp[\ -\ (te^{-X\hat{\beta}})^{\hat{\delta}}\ ]$$

and for proportional hazard model as

$$\hat{S}(t;X) = [\ \hat{S}_0(t)\ ]^{e^{X\hat{\beta}}}$$

After we calculated $\hat{S}(t;X)$ for both models, then we can calculate the probability of failure for each company during specific time t from $1 - \hat{S}(t;X) = \hat{F}$. If $F_c = 0.5$ is used as a cutoff probability value for classification (Collins and Green, 1982 and Ko, 1982), then companies are predicted to fail if this probability exceeds the critical level F_c and they are predicted not to fail if $\hat{F} < F_c$. Two types of errors are possible. Type I error is defined as predicting that a non-surviving company will survive and Type II error is defined as predicting a surviving company will fail.

The result of testing both models (parametric and nonparametric) in the two cases a(1 and 2) and b for $F_c = 0.5$ are shown in Table 6.7 below.

It can seen from the Table that the nonparametric model (proportional hazard model) for case (b) gives slightly better correctly classified percentage as well as smaller Type I and Type II errors than the results for the corresponding case b of the parametric model (Weibull regression model) and both models give better results for case b than case a. Therefore, the nonparametric model seems to produce the best fit for prediction purposes. Also, this model should be used in conjunction with the likelihood procedure for the classification of unknown companies given in Chapter 4.

Table 6.7 Test results of forecasting from parametric (Weibull regression model) and nonparametric (proportional hazard model) models of company failures for case a(1 and 2) (*before* reclassifying "merged" and "other" companies) and case b (*after* reclassifying "merged" and "other" companies).

Type of model	Parametric(Weibull regression) model			Nonparametric(proportional hazard) model		
	case a		case b	case a		case b
	(1)	(2)		(1)	(2)	
Type I error	13.1%	14.7%	8.2%	9.5%	14.3%	5.3%
Type II error	15.6%	17%	11.2%	14.8%	16.5%	10.7%
correctly classified percentage	84.4%	83.4%	89.2%	85.5%	83.9%	90.1%

CHAPTER SEVEN

SUMMARY AND CONCLUSIONS

This study consists of a statistical analysis of accounting data of British industrial companies in order to assess the forecasting ability of existing and new techniques, using financial ratios of large UK companies obtained from the EXSTAT source. Some methodological issues have been investigated and classification methods to discriminate between non-surviving and surviving companies have been revised accordingly.

Methodological issues

The methodological issues investigated are :

1- Imbalance between non-survivor and survivor groups.

As has been stated in Chapters 3 and 4, many existing failure studies were carried out using two groups of companies that have been matched on the basis of similarity in size, industrial classification and year of data. The sample size of the two groups (one non-surviving, one surviving) has usually been equal. However, the matching technique may not be correct because it does not give a fair representation of reality. Matched samples constructed with equal numbers of non-surviving and surviving companies cause both parameter and probability estimates to be asymptotically biased. This use of non-random, choice-based, equal-share samples in model estimation is a major criticism of many existing failure prediction models.

2- <u>Company failure and other forms of termination of company life</u>.

As a measure of financial distress, bankruptcy seems to be relatively objective if not comprehensive. However, the limited number of bankruptcies amongst companies with a record of adequate financial disclosure has led to a situation where sample sizes are small. For example , Altman (1968), Deakin (1972) and Taffler (1982) relied on samples of bankrupt companies numbering only 33, 32 and 23 respectively.

In his study of the failure of industrial enterprises quoted on the London Stock Exchange, Taffler defines company failure as "receivership, voluntary liquidation (creditors), winding up by court order or equivalent". However, the definition of failure has varied from study to study, and has been broadened to include various states of financial distress. For example, Altman (1968) and Ohlson (1980) restricted their sample to companies filing for bankruptcy, whereas Beaver (1966) and Blum (1974) included companies unable to pay their financial obligations within the "failed" group.

On the other hand, with regard to companies which cease trading, the company's life does not necessarily end with the liquidation of its remaining assets in order to pay off creditors in the context of receivership. There are various other ways of leaving a given population, such as delisting or transfer of residence, and a considerable number of companies which cease trading are taken over by, or merge with, other going concerns.

Companies which are taken over present a particular problem. For instance, rather than going into voluntary liquidation and selling its assets to an interested acquirer, the ownership of the entire firm may be transferred from the original shareholders to the acquiring company. Thus, a company which merges or is taken over may possibly have been heading for technical bankruptcy. On the other hand, the potential of a company which is bought out may well be reflected in a sound financial situation. A weakness of many previous studies has been the failure to incorporate such events into the probabilistic framework that underlies the outcomes of financial distress and survival.

3- Time dependent rate of failure.

Until recently (see Lau, 1987, for example), failure prediction research has been concerned primarily with single state models, e.g. discriminant and logit analyses, rather than with the probability of failure over a series of financial states. This does not show the probability of failure in a given time period or to allow us to estimate the conditional probability of a company failing between t and t+1, given that it had survived up to t (i.e. the hazard rate).

Solutions

1- The first issue i.e. the imbalance between non-survivor and survivor groups, is resolved through the use of

Randomly-censored stratified samples.

The approach adopted here is to apply stratified sampling in order to select randomly the survivors such that the failures in any given year when expressed as a proportion of the total number of non-survivors are reflected in the censored group but which nevertheless are matched to the lifetimes of the non-survivors. This solution is adopted because further problems arise in moving from a matched sampling basis, as the structure of the survivor group no longer reflects that of the non-survivors. For instance, in our study, the survivor group contains all companies which have not yet failed (i.e. in this data set, 337 usable survivors) and there are 95 non-survivors, of which 21 were bankrupts. Given the small number of listed companies failing each year, the need to generate a sufficiently large sample of non-survivors by including companies which failed at different points in time produces a problem in structuring the survivor group which is not met in the previous "matched sample" based studies as discussed in Chapter 4. The randomly-censored stratified samples procedure is therefore used to overcome this problem.

2- We resolved the second issue i.e. company failure and other forms of termination of company life, using a

<u>Likelihood estimation of failure for takeover targets</u>.

In order to estimate the likelihood that a company that has been acquired (or wound-up for reasons other than bankruptcy) posseses characteristics similar to those of failed companies or, alternatively, non-failed companies, we used the Weibull survival likelihood function to estimate the log-likelihood for the known bankrupts and for the survivors. Next, we add one observation first to the failed group and then to the survivor group, and recompute the log-likelihoods. A decrease in the log-likelihood indicates a worse fit of the model to the data, and we use this as our criteria for classifying takeover targets and others to the "non-survivor" and "survivor" groups.

3- The third issue i.e. time dependent rate of failure, is resolved using

Survival models.

Applying a new approach based on survival analysis, the following parametric models were used: exponential, Weibull and log-logistic regression models. The nonparametric proportional hazard model was also investigated. These models relate the hazard function to the length of survival time and the financial variables i.e. they take into account the length of company life. At the same time, these models give the estimate for the rate of failure and survival functions for the company. The overall adequacy of the models has been checked using residual analysis and we found that the Weibull regression model fitted the data better than the exponential and log-logistic regression models, and therefore classification tests for parametric models were restricted to the Weibull. The proportional hazard model also fitted the data adequately.

Conclusions and implications

(1) The results of multiple discriminant analysis using randomly-censored stratified sampling are given in Chapter 4, Tables 4.3b, 4.4b and 4.8b. Whilst the results confirm that the application of discriminant analysis to unbalanced samples provides weak results, it is noticeable that there is discriminatory power one year prior to failure, in spite of the severe imbalance between survivors and non-survivors. The procedure was repeated 32 times, and applied to (i) all non-survivors, (ii) bankrupt companies only, and (iii) the reclassified grouping where companies which were taken over

but which had a high likelihood of failure were reassigned along with the bankrupt companies to the "failures" group. The type I error falls from (i) 41.8% to (ii) 24.5% and (iii) 28.4% respectively, on average over the 32 runs.

This weak result may be largely due to the fact that the sample sizes for the two groups of companies are quite different as this is a major shortcoming of conventional modelling procedures, where sample bias is a feature of the data. Nevertheless, as is demonstrated in Chapter 4, Tables 4.3(a and b), 4.4(a and b) and 4.8(a and b), the application of randomly-censored stratified samples allows us to observe a substantial decrease in type I error one year before failure.

2- The use of the Weibull survival function resulted in 40 companies which had been acquired or wound-up for reasons other than bankruptcy being reclassified as "failures" and 34 non-failures which could be grouped with the surviving companies. This use of the Weibull survival procedure to reclassify those non-survivors which ceased trading without going into bankruptcy has the effect of decreasing type I and type II errors when using discriminant , logit and survival analyses as demonstrated in Chapters 4, 5, and 6.

3- These first results from the application of survival analysis provide a reasonable interpretation of the financial events leading to bankruptcy and indicate success in model estimation. For example, they infer that high Funds Flow to Current Liabilities and high Liquid Assets to Current liabilities are consistent with survival. Furthermore they provide a probabilistic estimate of the rate of failure and, for each of the parametric models and the proportional hazards model, there is a gradual increase in hazard with survival time followed by a decrease, which is consistent with existing assumptions about the bankruptcy process in the real world (as discussed in Chapter 6). The results shown in the following table, particularly with respect to the reduction in type I error, indicate the considerable potential of survival modelling of financial distress.

A comparison of linear discriminant analysis (LD), quadratic discriminant analysis (QD), logit, proportional hazards (PH) and the Weibull survival model

Cases	Type of model	LD model	QD model	Logit model	PH model	Weibull model
Before reclassifying takeovers, mergers and other terminations						
failed = bankrupt non-failed = survivors	Type I error	15.5%	38.1%	14.8%	9.5%	13.1%
	Type II error	18.0%	25.1%	16.8%	14.8%	15.6%
	Percentage correctly classified	82.3%	74.1%	83.4%	85.5%	84.4%
failed = bankrupt+others non-failed = survivors	Type I error	19.5%	33.7%	16.2%	14.3%	14.7%
	Type II error	22.3%	34.1%	17.9%	16.5%	17.0%
	Percentage correctly classified	78.2%	64.9%	82.6%	83.9%	83.4%
After reclassification						
failed = bankrupt + 40 others non-failed =survivors + 34 others	Type I error	13.5%	29.5%	11.5%	5.3%	8.2%
	Type II error	12.8%	32.9%	11.2%	10.7%	11.2%
	Percentage correctly classified	87.2%	67.6%	88.7%	90.1%	89.2%

As shown in the above table these results seem to suggest that analysis using survival models will produce a more accurate solution than linear discriminant analysis, quadratic discriminant analysis or logit model. Furthermore the proportional hazards model is a more promising approach for developing practical models for predicting financial distress.

In conclusion, this thesis has demonstrated the statistical power of survival modelling when applied to bankruptcy prediction, and has set down solutions to some of the methodological issues which are met when moving from the conventional techniques which have dominated the field to date.

APPENDIX 1 EXSTAT

EXSTAT is a service of company data in computer readable form. It covers over two thousand British, other European, Australian and Japanese quoted and unquoted concerns. The record for each company is comprised of four sections:

(a) section A - contains EXTEL control data and is available on card forms.

(b) section B, C (and data appendices CA, CB, CC, CD, CE for U.K./ Eire companies) and D, appear on customer tapes - and is where this study obtained its data.section B contains standing information such as country of registration and industrial classification. This section is also known as the company data section.

(c) section C contains accounts data.

(d) section D contains security data relevant to the accounts.

Therefore, for each company, there is always only one section B representing the standing information. For each year that data is held, there is a section C (and data appendices CA, CB, CC, CD, CE for U.K./ Eiro companies) containing balance sheets and profit and loss account items for the year. Also, a section D appears for any equity and preference shares in issue at the end of each year, giving the security data relevant to the accounts.

APPENDIX 2

List of non-surviving companies

NO.	Name of company
1-	ACROW PLC
2-	AERO NEEDLES GROUP PLC
3-	ALLEN(W.G.)& SONS(TIPTON)PLC
4-	AMALGAMATED INDUSTRIALS
5-	AMALGAMATED POWER ENGINEERING
6-	AVERYS
7-	BADGER LTD
8-	BAILEY(N.G.)& CO LTD
9-	BAMFORDS
10-	BEYER PEACOCK & CO
11-	BLAKEY'S(MALLEABLE CASTINGS)
12-	BRITISH ALUMINIUM CO PLC
13-	BRITISH ROLLMAKERS CORPN LTD
14-	BROCKHOUSE PLC
15-	BROTHERHOOD(PETER) PLC
16-	BROWN(DAVID)GEAR INDUSTRIES LTD
17-	BUCYRUS(U.K.)LTD
18-	CAPPER NEILL PLC
19-	CARBORUNDUM CO LTD
20-	CENTRAL MFG & TRADING GROUP LTD
21-	CENTRAL WAGON CO
22-	CLARKE CHAPMAN
23-	CLIFFORD(CHARLES)INDUSTRIES
24-	COLTNESS GROUP
25-	CORNERCROFT
26-	CRANE'S SCREW(HLDGS)
27-	DANKS GOWERTON PLC
28-	DERRITRON PLC
29-	DORMAN SMITH HLDGS
30-	DRAKE & SCULL HOLDINGS PLC
31-	DUCTILE STEELS
32-	DUNFORD & ELLIOTT
33-	ELKEM LTD

34- ELSWICK-HOPPER PLC

35- EMMS(THEODORE)

36- ENGLISH CARD CLOTHING CO

37- FAIRBAIRN LAWSON

38- FAIREY CO

39- FLUIDRIVE ENGINEERING CO

40- GALLENKAMP(A.)& CO

41- GENERAL ENGINEERING CO(RADCLIFFE)

42- GOSFORTH INDUSTRIAL HOLDINGS

43- GRAHAM WOOD STEEL GROUP

44- HALL-THERMOTANK LTD

45- HARTLE MACHINERY INTERNATIONAL

46- HAWKINS & TIPSON PLC

47- HAYTERS PLC

48- HEAD WRIGHTSON & CO

49- HUNT & MOSCROP GROUP PLC

50- INTERNATIONAL COMBUSTION(HLDGS)

51- JEAVONS ENGINEERING PLC

52- KLEEMAN INDUSTRIAL HLDGS

53- LE BAS(EDWARD)

54- LEADENHALL STERLING

55- LIGHTING & LEISURE INDUSTRIES

56- LINER CONCRETE MACHINERY CO

57- MARTIN(TOM)METALS GROUP

58- MOSS ENGINEERING GROUP

59- MUNFORD & WHITE PLC

60- NEGRETTI & ZAMBRA

61- NEWALL MACHINE TOOL CO

62- NEWMANS TUBES

63- NORMAND ELECTRICAL HOLDINGS

64- PACTROL ELECTRONICS PLC

65- PROTIMETER PLC

66- RCF HLDGS

67- RANK PRECISION INDUSTRIES(HLDGS)PLC

68- RECORD RIDGWAY

69- REDMAN HEENAN INTERNATIONAL PLC

70- REYROLLE PARSONS

71- ROBB CALEDON SHIPBUILDERS

72- SANDERSON KAYSER

73- SANGAMO WESTON LTD

74- SCOTT(JAMES)ENGINEERING GROUP

75- SERCK PLC

76- SHAW(FRANCIS)PLC

77- SHEFFIELD TWIST DRILL & STEEL CO.

78- SIG DAVALL PLC

79- SILVERTHORNE GROUP PLC

80- SIMPLEX (POWER CENTRE) LTD

81- SOLUS GROUP

82- SPOONER INDUSTRIES

83- STAR ALUMINIUM PLC

84- STONE-PLATT INDUSTRIES PLC

85- SWAN HUNTER GROUP

86- ULTRA ELECTRONIC HLDGS

87- UNITED ELECTRONIC HOLDINGS PLC

88- VOSPER PLC

89- WESTFORTH ELECTRICAL & AUTOMATION

90- WESTINGHOUSE BRAKE & SIGNAL CO

91- WILJAY PLC

92- WILSHAW SECURITIES PLC

93- WOLF ELECTRIC TOOLS(HLDGS)

94- WOLVERHAMPTON DIE CASTING GROUP

95- YOUNG, AUSTEN & YOUNG

APPENDIX 3

List of surviving companies

NO.	Name of company
1-	AEG(UK)LTD
2-	AIM GROUP PLC
3-	APV HOLDINGS PLC
4-	ASD PLC
5-	ADWEST GROUP PLC
6-	AEROSPACE ENGINEERING PLC
7-	ALCOA OF GREAT BRITAIN LTD
8-	ALFA-LAVAL CO LTD
9-	ALLIED HOLDINGS(U.K.)LTD
10-	ALUSUISSE(U.K.)LTD
11-	AMALGAMATED METAL CORP PLC
12-	AMARI PLC
13-	ANDERSON STRATHCLYDE PLC
14-	ANGLESEY ALUMINIUM LTD
15-	ANGLO NORDIC HOLDINGS PLC
16-	ARCOLECTRIC(HOLDINGS)PLC
17-	ARIEL INDUSTRIES PLC
18-	ARLEN PLC
19-	ASH & LACY PLC
20-	ASSOCIATED ELECTRICAL INDS LTD
21-	ASSOCIATED ENERGY SERVICES PLC
22-	ASTRA INDUSTRIAL GROUP PLC
23-	AURORA PLC
24-	AYRSHIRE METAL PRODUCTS PLC
25-	BICC PLC
26-	BM GROUP PLC
27-	BUSM CO LTD
28-	BABCOCK INDUSTRIAL ELEC.PRODS LTD
29-	BABCOCK INTERNATIONAL PLC
30-	BAILEY(C.H.)PLC
31-	BAILEY(N.G.)ORGANISATION LTD(THE)
32-	BAKER PERKINS PLC

33- BAMFORD(J.C.)EXCAVATORS LTD

34- BANRO INDUSTRIES PLC

35- BARDSEY PLC

36- BARR & STROUD LTD

37- BARTELLA LTD

38- BARTON GROUP PLC

39- BAXI PARTNERSHIP LTD

40- BEAUFORD GROUP PLC(THE)

41- BENFORD CONCRETE MACHINERY PLC

42- BESTOBELL PLC

43- BETEC PLC

44- BEVAN(D.F.)(HOLDINGS)PLC

45- BILLAM(J.)PLC

46- BIRMID QUALCAST PLC

47- BIRMINGHAM MINT GROUP PLC

48- BLACK & DECKER

49- BLACKETT HUTTON HOLDINGS LTD

50- BODYCOTE INTERNATIONAL PLC

51- BOGOD-PELEPAH PLC

52- BOOTHAM ENGINEERS PLC

53- BOSCH(ROBERT)LTD

54- BOULTON(WILLIAM)GROUP PLC

55- BOVING & CO LTD

56- BRAIME(T.F.& J.H.)(HOLDINGS)PLC

57- BRAMMER PLC

58- BRASWAY PLC

59- BRICKHOUSE DUDLEY PLC

60- BRIDON PLC

61- BRIDPORT-GUNDRY PLC

62- BRITANNIA REFINED METALS LTD

63- BRITISH AEROSPACE PLC

64- BRITISH ALCAN ALUMINIUM LTD

65- BRIT.MANUFACTURE & RESEARCH CO LTD

66- BRITISH SHIPBUILDERS

67- BRITISH STEAM SPECIALTIES GROUP PLC

68- BRITISH STEEL CORPORATION

69- BROMSGROVE INDUSTRIES PLC

70- BRONX ENGINEERING HOLDINGS PLC

71- BROOKE TOOL ENGINEERING(HLDGS)PLC

72- BROWN(DAVID)HOLDINGS LTD

73- BROWN(JOHN)PLC

74- BRUNTONS(MUSSELBURGH)PLC

75- BULGIN(A.F.)& CO PLC

76- BULLERS PLC

77- BULLOUGH PLC

78- BURGESS PRODUCTS(HOLDINGS)PLC

79- CEF HOLDINGS LTD

80- CI GROUP PLC

81- CAMBRIDGE INSTRUMENT CO PLC(THE)

82- CAMERON IRON WORKS LTD

83- CARCLO ENGINEERING GROUP PLC

84- CASTINGS PLC

85- CATERPILLAR TRACTOR CO LTD

86- CELTIC HAVEN PLC

87- CHAMBERLIN & HILL PLC

88- CHEMRING GROUP PLC

89- CHICAGO PNEUMATIC HOLDINGS LTD

90- CHLORIDE GROUP PLC

91- CHRISTY HUNT PLC

92- CINCINNATI MILACRON LTD

93- CIRCAPRINT HOLDINGS PLC

94- CLARKE(T.)PLC

95- CLAYTON,SON & CO(HOLDINGS)PLC

96- CLYDE BLOWERS PLC

97- COGHLANS PLC

98- COHEN(A.)& CO PLC

99- COMBINED ELECTRICAL MFRS LTD

100- COMPAIR LTD

101- CONCENTRIC PLC

102- COOK(WILLIAM)& SONS(SHEFFIELD)PLC

103- COOPER(FREDERICK)PLC

104- CRANE LTD

105- CRONITE GROUP PLC(THE)

106- CROWN HOUSE PLC

107- CUMMINS U.K.LTD

108- DALE ELECTRIC INTERNATIONAL PLC

109- DAVIES & METCALFE PLC

110- DAVY CORPORATION PLC

111- DELTA GROUP PLC

112- DENMANS ELECTRICAL PLC

113- DERITEND STAMPING PLC

114- DESOUTTER BROTHERS(HOLDINGS)PLC

115- DEWHURST PLC

116- DEXION-COMINO INTERNATIONAL LTD

117- DICKIE(JAMES)& CO(DROP FORGINGS)PLC

118- DOBSON PARK INDUSTRIES PLC

119- DOM HOLDINGS PLC

120- DOWDING & MILLS PLC

121- DOWNIEBRAE HOLDINGS PLC

122- DOWTY GROUP PLC

123- DURACELL BATTERIES LTD

124- DYSON(J.& J.)PLC

125- EIS GROUP PLC

126- ECOBRIC HOLDINGS PLC

127- EDMUNDSON ELECTRICAL LTD

128- ELBAR INDUSTRIAL PLC

129- ELBIEF PLC

130- ELLIOTT(B.)PLC

131- ELSWICK PLC

132- ENGELHARD LTD

133- EVA INDUSTRIES PLC

134- EVERED HOLDINGS PLC

135- FKI ELECTRICALS PLC

136- FR GROUP PLC

137- FARMER(S.W.)GROUP PLC

138- FARREL BRIDGE LTD

139- FENNER(J.H.)(HOLDINGS)PLC

140- FIFE INDMAR PLC

141- FIRTH(G.M.)(HOLDINGS)PLC

142- FLEXELLO CASTORS & WHEELS PLC

143- FLUOR(GREAT BRITAIN)LTD

144- FOLKES GROUP PLC

145- FOSTER WHEELER LTD

146- G.E.I.INTERNATIONAL PLC

147- GARTON ENGINEERING PLC

148- GLYNWED INTERNATIONAL PLC

149- GODWIN WARREN CONTROL SYSTEMS PLC

150- GOODWIN PLC

151- GORING KERR PLC

152- GOTAVERKEN ARENDAL INVESTMENTS LTD

153- GREENBANK GROUP PLC

154- GROSVENOR GROUP PLC

155- HABIT PRECISION ENGINEERING PLC

156- HADEN LTD

157- HALL ENGINEERING(HOLDINGS)PLC

158- HALL(MATTHEW)PLC

159- HALLIBURTON MFG & SERVICES LTD

160- HALMA PLC

161- HAMPSON INDUSTRIES PLC

162- HARRIS(PHILIP)(HOLDINGS)PLC

163- HAWKER SIDDELEY GROUP PLC

164- HAY(NORMAN)PLC

165- HEATH(SAMUEL)& SONS PLC

166- HEWITT(J.)& SON(FENTON)PLC

167- HILL & SMITH HOLDINGS PLC

168- HOPKINSONS HOLDINGS PLC

169- HOWDEN GROUP PLC

170- HUMBERSIDE ELECTRONIC CONTROLS PLC

171- HUMPHREYS & GLASGOW LTD

172- HUNSLET(HOLDINGS)PLC

173- HUNTING ASSOCIATED INDUSTRIES PLC

174- HYSTER LTD

175- IMI PLC

176- INCO EUROPE LTD
177- INCO ENGINEERED PRODUCTS LTD
178- INGERSOLL-RAND CO LTD
179- INGERSOLL-RAND HOLDINGS LTD
180- INTERNATIONAL MILITARY SERVICES LTD
181- JACKSON(J.& H.B.)PLC
182- JOHNSON & FIRTH BROWN PLC
183- JOHNSON MATTHEY PLC
184- JONES & SHIPMAN PLC
185- JONES STROUD(HOLDINGS)PLC
186- KAISER INTERNATIONAL(U.K.)LTD
187- KAYE ORGANISATION LTD(THE)
188- LPA INDUSTRIES PLC
189- LAIRD GROUP PLC
190- LANCERBOSS GROUP LTD
191- LAURENCE SCOTT LTD
192- LEE(ARTHUR)& SONS PLC
193- LEE(ARTHUR)& SONS(HOT ROL.MILLS)PLC
194- LEXA LTD
195- LEY'S FOUNDRIES & ENGINEERING PLC
196- LILLESHALL COMPANY PLC(THE)
197- LINREAD PLC
198- LLOYD(F.H.)HOLDINGS PLC
199- LOCKER(THOMAS)(HOLDINGS)PLC
200- LONDON & SCAND.METALLURGICAL CO LTD
201- M.K.ELECTRIC GROUP PLC
202- M.L.HOLDINGS PLC
203- MS INTERNATIONAL PLC
204- MCARTHUR GROUP LTD
205- MCKECHNIE BROTHERS PLC
206- MACLELLAN(P.& W.)PLC
207- MANGANESE BRONZE HOLDINGS PLC
208- MARSHALL(THOMAS)(LOXLEY)PLC
209- MARTIN-BAKER AIRCRAFT CO LTD
210- MARTIN-BAKER(ENGINEERING)LTD
211- MARTONAIR INTERNATIONAL PLC
212- MASSEY-FERGUSON HOLDINGS LTD

213- MATHER & PLATT LTD
214- MEGGITT HOLDINGS PLC
215- METALRAX GROUP PLC
216- MITCHELL COTTS PLC
217- MITCHELL SOMERS PLC
218- MOLINS PLC
219- MOLYNX HOLDINGS PLC
220- MOTHERCAT LTD
221- MYSON GROUP PLC
222- NATIONAL SUPPLY CO(U.K.)LTD
223- NEEPSEND PLC
224- NEIL & SPENCER HOLDINGS PLC
225- NEILL(JAMES)HOLDINGS PLC
226- NEWEY GROUP LTD
227- NEWMARK(LOUIS)PLC
228- NEWTON,CHAMBERS & COMPANY PLC
229- NOBLE & LUND PLC
230- NORANDA SALES CORPN OF CANADA LTD
231- NORTH BRITISH STEEL GROUP(HLDGS)PLC
232- NORTHERN ENGINEERING INDUSTRIES PLC
233- NORTON ABRASIVES LTD
234- OLDHAM BATTERIES LTD
235- OTIS ELEVATOR PLC
236- OXFORD INSTRUMENTS GROUP PLC(THE)
237- PCT GROUP PLC
238- PALSA HOLDINGS LTD
239- PARKER(FREDERICK)GROUP PLC
240- PARKFIELD GROUP PLC
241- PARSONS(RALPH M.)CO LTD(THE)
242- PEGLER-HATTERSLEY PLC
243- PENNWALT HOLDINGS LTD
244- PERKIN-ELMER LTD
245- PETBOW HOLDINGS PLC
246- PLASTIC CONSTRUCTIONS PLC
247- PLESSEY CONNECTORS LTD
248- POLYMARK INTERNATIONAL PLC

249- PORTALS HOLDINGS PLC
250- PORTER CHADBURN PLC
251- PRATT(F.)ENGINEERING CORPN PLC
252- PRESS TOOLS PLC
253- PRIEST(BENJAMIN)GROUP PLC
254- PROCON(GREAT BRITAIN)LTD
255- RHP GROUP PLC
256- RTD GROUP PLC
257- RADIANT METAL FINISHING PLC
258- RAMCO OIL SERVICES PLC
259- RANSOMES SIMS & JEFFERIES PLC
260- RATCLIFFE(F.S.)INDUSTRIES PLC
261- RATCLIFFS(GREAT BRIDGE)PLC
262- RENISHAW PLC
263- RENOLD PLC
264- RESTMOR GROUP PLC
265- RICARDO CONSULTING ENGINEERS PLC
266- RICHARDS(LEICESTER)PLC
267- RICHARDSONS,WESTGARTH PLC
268- ROBINSON(THOMAS)GROUP PLC
269- ROCKWELL INTERNATIONAL LTD
270- ROLLS-ROYCE PLC
271- ROSSER & RUSSELL LTD
272- ROTAFLEX PLC
273- ROTORK PLC
274- RUSTON & HORNSBY LTD
275- S.I.GROUP PLC
276- SKF(UK)LTD
277- STC DISTRIBUTORS LTD
278- SANDVIK LTD
279- SAVILLE GORDON(J.)GROUP PLC
280- SCHLUMBERGER MEASUREMT&CONTL(UK)LTD
281- SCHOLES(GEORGE H.)PLC
282- SENIOR ENGINEERING GROUP PLC
283- SEVALCO LTD
284- SHEERNESS STEEL CO PLC

285- SHEFFIELD INSULATING CO LTD(THE)

286- SHEFFIELD SMELTING CO LTD

287- SIEMENS LTD

288- SIMON ENGINEERING PLC

289- SIMPLEX ELECTRICAL HOLDINGS LTD

290- 600 GROUP PLC

291- SLINGSBY(H.C.)PLC

292- SMITH WHITWORTH PLC

293- SMITHS INDUSTRIES PLC

294- SOUND DIFFUSION PLC

295- SPEAR & JACKSON INTERNATIONAL PLC

296- SPENCER CLARK METAL INDUSTRIES PLC

297- SPIRAX-SARCO ENGINEERING PLC

298- STAINLESS METALCRAFT PLC

299- STANELCO PLC

300- STANLEY WORKS LTD(THE)

301- STERLING INDUSTRIES PLC

302- STONE INTERNATIONAL PLC

303- STOTHERT & PITT PLC

304- SULZER BROS(UK)LTD

305- SUMNER PRODUCTS PLC

306- SYCAMORE HOLDINGS PLC

307- SYKES(HENRY)PLC

308- SYLTONE PLC

309- SYMONDS ENGINEERING PLC

310- T.I.GROUP PLC

311- TSL THERMAL SYNDICATE PLC

312- TACE PLC

313- TECHNOLOGY INCORPORATED(UK)LTD

314- TEKTRONIX U.K.LTD

315- TELFOS HOLDINGS PLC

316- TEX HOLDINGS PLC

317- TEXTRON LTD

318- THORPE(F.W.)PLC

319- TRIEFUS PLC

320- TRIPLEX PLC

321- TWIL LTD
322- TYZACK(W.A.)PLC
323- TYZACK TURNER PLC
324- UNITED SCIENTIFIC HOLDINGS PLC
325- UNITED SPRING & STEEL GROUP PLC
326- UNITED WIRE GROUP PLC
327- VICKERS PLC
328- VICTOR PRODUCTS PLC
329- VINTEN GROUP PLC
330- VOLEX GROUP PLC
331- WA HOLDINGS PLC
332- WPP GROUP PLC
333- WADKIN PLC
334- WAGON INDUSTRIAL HOLDINGS PLC
335- WALKER(C.)& SONS LTD
336- WALKER(C.& W.)HOLDINGS PLC
337- WALKER,CROSWELLER & CO LTD
338- WALKER(THOMAS)PLC
339- WATSHAM'S PLC
340- WATSON(R.KELVIN)PLC
341- WEIR GROUP PLC
342- WELLMAN PLC
343- WEST BROMWICH SPRING PLC(THE)
344- WESTERN SELECTION PLC
345- WESTLAND PLC
346- WESTWOOD DAWES PLC
347- WHESSOE PLC
348- WHEWAY PLC
349- WHITTINGTON ENGINEERING COMPANY P
350- WHITWORTH ELECTRIC(HOLDINGS)PLC

REFERENCES

Altman, E. (1968). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", Journal of Finance, Vol. 23, No. 4, PP. 589 - 609.

Altman, E. (1971). "Corporate bankruptcy in America", Lexington, mass., Heath-Lexington, Co..

Altman, E. , Haldeman, P. and Narayanan, P. (1977). "ZETA analysis. A new model to identify bankruptcy risk of corporations", Journal of Banking and Finance, June, pp. 29-54.

Altman, E. (1984). "The success of business failure prediction models - an international survey", Journal of Banking and Finance, Vol. 8, PP. 171 - 198.

Aalen, O. O. (1989). "A linear regression model for the analysis of lifetimes", Statistics in Medicine, Vol. 8, PP. 907 - 925.

Aitkin, M. and Clayton, D. (1980). "The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM", Applied Statistics, 29, No. 2, PP. 156 - 163.

Aitkin, M., Laird, N. and Francis, B. (1983). "A reanalysis of the stanford heart transplant data", Journal of American Statistics Association, Vol. 78, No. 382, PP. 264 - 292.

Aitkin, M., Anderson, D., Francis, B. and Hinde, J. (1989). "Statistical modeling in GLIM", Guildfor & King's Lynn.

Al-moswie, A. M. (1982). "Distribution-free methods of discriminant", MSC thesis, university of wales, swansea.

Andersen, P. K. and Gill, R. D. (1982). "Cox's regression model for counting processes: A large sample study", Annals of Statistics , Vol. 10, No. 4, PP. 1100 - 1120.

Andersen, P. K., Johnsen, K. B., Deckert, T. ,Green, A., Hougaard, P., Keiding, N. and Keriner, S. (1985). "A Cox regression model for the relative mortality and its application to diabetes mellitus survival data", Biometrics, 41, PP. 921 - 932.

Barlow, R. E. and Proscham, F. (1975). "Statistical theory of reliability and life testing: Probability models", Holt, Rinehart and Winston, Inc..

Barnes, P. (1982). "Methodological implications of non-normally distributed financial ratios", Journal of Business Finance and Accounting 9, 1, pp. 51-61.

Barron, M. (1986). "Year-end heterogeneity in calculations of industru and economy averages of accounting numbers", Accounting and Business Research, Autumn, pp. 275 - 284.

Barnes, p. (1990). "The prediction of takeover Targets in the U.K. by means of multiple discriminant analysis", Journal of Business Finance and Accounting, 17, (1), pp. 73-84.

Beaver, W. H. (1966). "Financial ratios as predictors failure", Journal of Accounting Research, Vol. 5, PP. 71 - 111.

Bedingfield, J. P., Peckers, P. M. J. and Stagliano, A.J. (1985). "Distributions of financial ratios in the commercial banking industry", Journal of Financial Research, vol. 8, no. 1, pp. 77-81.

Berkman, J. D., Brownstons, and Associates (1979). "Quail 4.0 Users' Manual", Berkeley : Department of Economics, University of California.

Betts, J and Belhoul, D. (1982). "The identification of companies in danger of failure using discriminant analysis", proceedings of 7th symposium on advances in reliability technology, April, Held Bradford university.

Blum, M. P. (1974). "Failing company discriminant analysis", Journal of Accounting Research, Spring, PP. 1 - 25.

Bougen, P. D. and Drury, J. C. (1980). "U.K. statistical distributions of financial ratios, 1975", Journal of Business Finance and Accounting, Spring, Vol. 7, No.

1, PP. 39 - 47.

Breslow, N. (1974). "Covariance analysis of censored survival data", Biometrics, 30, PP. 89 - 99.

Burridge, J. (1982). "Statistical analysis of grouped lifetime data in a changing environment", ph.D. thesis, university of London.

Chen, K. H. and Shimerda, T. A. (1981). " An empirical analysis of useful financial ratios", Journal of Financial Management, Spring, PP. 51 - 60.

Conover, W. J. (1971). "Practical nonparametric statistics", John Wiley & Sons Inc., New York.

Cox, D. R. (1972). "Regression models and life tables (with discussion), Journal of Royal Statistical Society, B, 34, PP. 187 - 220.

Cox, D. R. (1975). "Partial likelihood", Biometrika, 62, PP. 269 - 276.

Cox, D. R. and Oakes, D. (1984). "Analysis of survival data", Chapman and Hall ltd., London.

Collins, R. A. and Green, R. D. (1982). "Statistical methods for bankruptcy forecasting", Journal of Economics and Business, March, PP. 349 - 354.

Crapp, H. R. and Stevenson, M. (1987). "Development of a method to assess the relevant variables and the probability of financial distress", Australian Journal of Management, Vol. 12, No. 2, PP. 221 - 235.

Dambolena, I. G. and Khoury, S. J. (1980). "Ratio stability and corporate failure", Journal of Finance, Vol. 35, No. 4, PP. 1017-1026.

Davis, P. T. (1985). "Competing risk survival analysis with applications to tyre testing", MSC thesis, university of Birmingham.

Dewing, A. S. (1941). "The financial policy of corporations", Ronald Press, New York.

Deakin, E. (1972). "A discriminant analysis of predictors of business failures", Journal of Accounting Research, Spring, PP. 167 - 179.

Deakin, E. B. (1976). "Distributions of financial accounting ratios: Some empirical evidence", Journal of Accounting Review, January, PP. 90 - 96.

Deakin, E. B. (1977). "Business failure prediction: An empirical analysis", Financial crises: Institutions and Markets in a Fragile Environment, Edited by Altman, E. and Sametz, A., John Wiley & Sons.

Draper, N. R. and Smith, H. (1981). "Applied regression analysis", second edition, John Wiley & Sons, Inc., London.

Dun and Brasdstreet (1982). "Key Business ratios", New York.

Elam, R. (1975). "The effect of lease data on the predictive ability of financial ratios", Journal of Accounting Review, January, PP. 25 - 43.

Edmister, R. (1972). "An empirical test of financial ratios analysis for small Business failure prediction", Journal of Financial and Quantitative Analysis, march.

Eisenbeis, R. A (1977). "Pitfalls in the application of discriminant analysis in Business, Finance, and Economics", Journal of Finance, Vol. 32 No. 3, PP. 875-900.

Foster, G. (1986). "Financial statement analysis", second edition, Prentice-Hall international, Englewood Cliffs, New Jersey.

Gentry, J. A., Newbold, P. and Whitford, D. J. (1985). "Classification bankrupt firms with funds flow components", Journal of Accounting Research, Vol. 23, No. 1, PP. 146 - 160.

Gordon, M. J. (1971). "Towards a theory of financial distress", Journal of Finance, Vol. 26, PP. 347 - 362.

Gombola, M. J. and Ketz, J. E. (1983). " Financial ratio patterns in retail and manufacturing organizations", Journal of Financial Management, Summer, PP. 45 - 56.

Gross, A. J. and Clark, V. A. (1975). "Survival distributions: Reliability applications in the biomedical sciences", Wiley, New York.

Hamer, M. M. (1983). "Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets", Journal of Accounting and Public Policy, 2, PP. 289 - 307.

Hopwood, W., Mckeown, J., Mutchler, J. (1988). "The sensitivity of financial distress prediction models to departures from normality", Contemporary Accounting Research, Vol. 5, No. 1, pp. 284 - 298.

Ingram, D. D. and Kleinman, J. C. (1989). "Empirical comparisons of proportional hazards and logistic regression models", Statistics in Medicine, Vol. 8, PP. 525 - 538.

Jeffers, J. P. (1978). "An introduction to system analysis : With ecological applications", William Clowes and Sons ltd, London.

Johnson, R. C. and Johnson, N. L. (1980). "Survival models and data analysis", Wiley, New York.

Johnson, R. A. and Wichern, D. W. (1982). "Applied multivariate statistical analysis", Prentice-Hall, Inc., Englewood Cliffs, New York.

Kaplan, E. L. and Meier, P. (1958). "Nonparametric estimation from incomplete observation", Journal of American Statistical Association, June, PP. 458 - 481.

Kalbfleisch, J. D. and Prentice, R. L. (1973). "Marginal likelihoods based on Cox's regression and life model", Biometrika, 60, PP. 267 - 275.

Kalbfleisch, J. D. and Prentice, R. L. (1980). "The statistical analysis of failure time data", Wiley, New York.

Kay, R. (1977). "Proportional hazard regression models and the analysis of censored survival data", Applied Statistics, 26, No. 3, PP. 227 - 237.

Kendall, M. (1980). "Multivariate analysis", second edition, Charls Griffin and Company ltd., London.

Keasey, K. and Watson, R. (1987). "Non-financial symptoms and the prediction of small company failure: A test of argenti's hypotheses", Journal of Business Finance and Accounting, 14(3), PP. 335 - 354.

Ko, C. J. (1982). "A delineation of coporate appraisal models and classification of bankruptcy firms in Japan, ph.D.

thesis, New York university, New York.

Krall, J. M., Uthoff, V. A. and Harley, J. B. (1975). "A step-up procedure for selecting variables associated with survival", Biometrics, 31, PP. 49 - 57.

Krzanowski, W. J. (1977). "The performance of Fisher's linear discriminant under non-optimal conditions", Technometrics, 19, No. 2, PP. 191 - 200.

Lacbenbruch, P. A. (1975). "Discriminant analysis", Macmillan publishing Co., Inc., New York.

Lancaster ,T. and Nichell, S. (1980). "The analysis of re-employment probabilities for the unemployed (with discussion)", Journal of Royal Statistical Society, A, 143, PP. 141 - 165.

Lagakos, S. W. (1980). "The graphical evaluation of explanatory variables in proportional hazard regression models", Biometrika, 68, 1, PP. 93 - 98.

Lau, A. H. (1982). "On the prediction of firms in financial distress, with an evaluation of alternative funds-flow concepts", PH.D. thesis, university of Washington.

Luk, T. W. (1984). "Corporate Failure - A multivariate analysis of U.K. Companies", MSC. thesis, University of Southampton, Faculty of Social Sciences.

Lau, A. H. (1987). "A five-state financial distress prediction model", Journal of Accounting Research, Vol. 25, No. 2, PP. 127 - 138.

Lawless, J. F. (1982). "Statistical models and methods for lifetime data", Wiley, New York.

Lawless, J. F. (1983). "Statistical methods in reliability", Technometrics, Vol. 25, No. 4, PP. 305 - 316.

Lev, B. (1974). "Financial statement analysis: A new approach", Englewood Cliffs, N. J., Prentice-Hall Inc.

Lee, E. T. (1980). "Statistical methods for survival data analysis", Belmont, CA: Lifetime learning publications.

Mann, N. R., Schafer, R. E. and Singpurwalla, N. D. (1974). "Methods for statistical analysis of reliability and life data", Wiley, New York.

Martin, L. G. (1975). "The 500: A report on two decades", Fortune, May, Vol. 91, No. 5, PP. 238 - 241.

Manson, R. J. and Harris, F. C. (1979). "Predicting company failure in the construction industry", proceedings of the institution of Civil Engineers, 66, Part 1, May, PP. 301 - 307.

Mardia, K. V., Kent, J. T. and Bibby, M. J. (1979). "Multivariate analysis", Academic Press Inc., ltd, London.

Manly, B. F. J. (1986). "Multivariate statistical methods A primer", Chapman and Hall, London.

Mccullagh, P. and Nelder, J. A. (1983). "Generalized linear models", Chapman and Hall ltd, USA.

Meyer, P. A. and Pifer, H. W. (1970). "Prediction of bank failures", Journal of Finance, PP. 853 - 868.

Miller, R. G., Gong, G. and Munoz, A. (1981). "Survival analysis", Wiley, New York.

Morrison, D. F. (1976). "Multivariate statistical methods", second edition, Mcgraw-Hill, Inc., USA.

Moreau, T., O'quigley, J. and Mesbah, M. (1985). "A global goodness-of-fit statistic for the proportional hazards model", Applied Statistics, 34, No. 3, PP. 212 - 218.

Nelson, W. (1972). "Theory and applications of hazard plotting for censored failure data", Technometrics, Vol. 4, No. 4, PP. 945 - 951.

Nelson, W. (1973). "Analysis of residuals from censored data", Technometrics, Vol. 15, No. 4, PP. 697 - 715.

Ohlson, J. A. (1980). "Financial ratios and the probabilistic prediction of bankruptcy", Journal of Accounting Research, Vol. 18, No. 1, PP. 109 - 131.

Prentice, R. L. and Gloeckler, L. A. (1978). "Regression analysis of grouped survival data with application to breast cancer", Biometrics, 34, PP. 57 - 67.

Palepu, K. G. (1986). "Predicting takeover targets: A methodological and empirical analysis", Journal of Accounting and Economics, 8, pp. 3 - 35.

Schoenfeld, D. (1980). "Chi-squared goodness-of-fit tests for the proportional hazards regression model", Biometrika, 67, PP. 145 - 153.

Seber, G. A. F. (1984). "Multivariate observations", John Wiley & Sons, New York.

Snedecor, G. W. and Cochran, W. G. (1980). "Statistical methods", seventh edition.

Srivastava, M. S. and Carter, E. M. (1983). "An introduction to applied multivariate statistics", Elsevier Science publishing Co., Inc., New York.

Shailer, G. E. P. (1990). "The multivariate performance of Alternative Accounting variables for predicting unlisted Company", British Accounting Review, 22, P. 151-161.

Tisshaw, H. J. (1976). "Evaluation of downside risk using Financial ratios", MSC. thesis, City University Business School, London.

Taffler, R. J. (1982). "Forecasting company failure in the U.K. using discriminant analysis and financial ratio data", Journal of Royal Statistical Association, A, 145, Part 3, PP. 342 - 358.

Turnbull, B.W., Brown, B. W. and Hu, M. (1974). "Survivorship analysis of heart transplant data", Journal of American Statistical Association, Vol. 69, No. 345, PP. 74 - 80.

Walter, J. E. (1957). "Determination of technical solvency", Journal of Business, Vol. 30, No. 1, PP. 30 - 37.

Zippix, C. and Armitage, P. (1966). "Use of concomitant variables and incomplete survival information in the estimation of an exponential survival parameter", Biometrics, December, PP. 665 - 672.

Zmijewski, M.E. (1984). "Methodological issues related to the estimation of financial distress prediction models", Journal of Accounting Research Supplement, pp. 59-82.

COMPUTING REFERENCES

EXSTAT User Manual. January, 1981.

MINITAB Reference Manual, releases 7, 1989, minitab Inc..

SAS Institute Inc. SUGI supplemental Library User's Guide version 5 Edition, SAS Institute Inc., cary, North Carolina, 1986.

SPSS-X User's Guide, 3rd Edition, 1988.