An Empirical Analysis of Takeover Predictions in the UK:

Application of Artificial Neural Networks and Logistic Regression

by

Asim Yuzbasioglu

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

University of Plymouth Business School

May 2002

ABSTRACT

Asim Yuzbasioglu

An Empirical Analysis of Takeover Predictions in the UK: Application of Artificial Neural Networks and Logistic Regression

This study undertakes an empirical analysis of takeover predictions in the UK. The objectives of this research are twofold. First, whether it is possible to predict or identify takeover targets before they receive any takeover bid. Second, to test whether it is possible to improve prediction outcome by extending firm specific characteristics such as corporate governance variables as well as employing a different technique that has started becoming an established analytical tool by its extensive application in corporate finance field.

In order to test the first objective, Logistic Regression (LR) and Artificial Neural Networks (ANNs) have been applied as modelling techniques for predicting target companies in the UK. Hence by applying ANNs in takeover predictions, their prediction ability in target classification is tested and results are compared to the LR results. For the second objective, in addition to the company financial variables, non-financial characteristics, corporate governance characteristics, of companies are employed. For the fist time, ANNs are applied to corporate governance variables in takeover prediction purposes. In the final section, two groups of variables are combined to test whether the previous outcomes of financial and non-financial variables could be improved.

However the results suggest that predicting takeovers, by employing publicly available information that is already reflected in the share price of the companies, is not likely at least by employing current techniques of LR and ANNs. These results are consistent with the semi-strong form of the efficient market hypothesis.

LIST OF CONTENTS **Page**

CHAPTER 1
INTRODUCTION

CHAPTER 2
MOTIVES FOR TAKEOVERS, AGENCY THEORY AND CONTROL ENVIRONMENT

CHAPTER 3
CLASSIFICATION METHODS IN TAKEOVER PREDICTION STUDIES

CHAPTER 5

CORPORATE GOVERNANCE DATA MODELLING AND RESULTS

CHAPTER 6

COMBINED MODELS (FINANCIAL AND CORPORATE GOVERNANCE) AND RESULTS

CHAPTER 7

CONCLUSIONS

LIST OF FIGURES **Page**

LIST OF TABLES **Page**

LIST OF DIAGRAMS **Page**

Acknowledgements

The completion of this thesis is due to the contributions and assistance of many individuals. At this point I would like to express my deepest thanks and gratitude to Professor John Pointon for his guidance, scholarly advice and encouragement and to Dr. Peter Atrill for his encouragement and support. I also owe many thanks to the members of the staff at the Business School; Zofjia Wills and Leslie Lislei for their continuos help.

Many friends helped me during my research and made my stay in Plymouth enjoyable and intellectually edifying, I thank all of them; Simon Down, Linda Down, Dave Smith, Richenda Sabine, Umit Ulgen, Tom Wilson, Sarah Claydon, Kadiatou Yansene, and I especially thank Dr. Yurt Alici for his constant support.

Finally, my greatest gratitude must go to my family, who stood by me throughout the years of my stay abroad. They have been a constant source of hope and support. Without their encouragement and sacrifice this work has never been completed.

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been register for any other University award.

The study was financed and carried out with the University of Plymouth Business School.

Signed..

Date...31..6.04..................................

Chapter 1 - Introduction

1.1 Introduction

Mergers and takeovers are used as a strategy for corporate control and expansion, market stability, and in certain cases to provide excess gains to the shareholders of acquired firms (Dietrich and Sorensen, 1984). The rationale behind mergers and takeovers can be related to a wide range of factors. These include capital market liquidity, strategic reasons, business cycles, the economic and political environment, integration of the European Union, management motives, diversification, defence against acquisitions, and taxation. (Post, 1994; Brealey and Myers, 1991; Sudarsanam, 1995; Mueller, 1989).

In general, the explanation of why firms become takeover targets was first given by Manne (1965, p.112) as poor management;

"As an existing company is poorly managed-in the sense of not making as great return for the shareholders as could be accomplished under the other feasible managements-the market price of the shares declines relative to the shares of other companiesin the same industry or relative to the market as a whole".

Clearly, as a result of this logic, one can postulate that managers of other firms would take the low price signalling of stock market as evidence of poor performance on behalf of incumbent management and act accordingly by bidding for these low priced firms on the market.

However, the assumption of low pricing/valuation of inefficient management by the stock market relies on explicit presumption of efficient markets, which means that market prices of securities will always equal the fair or fundamental values of those securities. On the other hand, if one observes that stock prices move randomly approximating a random walk, the stock market will, randomly, under price some of the securities. This may not necessarily be as a result of its managers or mismanagement, but leaves them vulnerable to any takeover bid from a competitor or any firm that is randomly enjoying a period of high share price performance. High share price will also enable firms to raise necessary finance for the cost of the takeovers (Scherer, 1988).

Empirical studies have been carried out to predict acquisition targets or profile target companies' characteristics- before a takeover bid - mainly used publicly available financial information as financial ratios. These studies include work by Simkowitz and Monroe, 1971; Stevens,1973; Singh, 1975; Belkaoui, 1978; Wansley and Lane, 1983; Dietrich and Sorensen, 1984; Rege, 1984; Walkling, 1985; Palepu, 1986; Barnes, 1990; Sen *et al,* 1995; Powell, 1997; and Barnes. Most of these studies have reported successful prediction models for takeovers.

However, as indicated by Palepu (1986), contrary to the reports of successful prediction modelling, the stock market, through share prices, is unable to predict target firms in reality even within a very short time window. Dodd and Ruback (1977), and Asquith (1983) found that the market only receives the takeover news shortly before the announcement dates. As the most frequently cited line in takeover studies says 'it is difficult, if not impossible, for the market to predict future targets' Jensen and Ruback (1983, p.29).

In these studies statistical classification techniques of Dichotomous Classification test (DC), Multiple Discriminant Analysis (MDA) and Logistic Regression (LR) are used as predictive tools in corporate takeover prediction. Although almost all of them claimed a high degree of prediction (from 60% to 90%) in their classification of companies as targets and non-targets, Palepu (1986) shows that most of these studies suffer from sampling biases. This bias occurs when the hold-out sample is selected to test the predictive ability of the model, in that the non-targets are underrepresented in the sample. Due to this they have been able to claim a high degree of success in their model's classification or prediction of corporate takeovers. Palepu (1986), by correcting this sampling bias, estimated that logistic regression is capable of predicting only 45% of the targets and non-targets in a hold-out sample. Thus, any abnormal returns for investors, by simply predicting takeovers or using these techniques, seem unlikely.

Prediction studies in corporate takeovers are founded/modelled on corporate failure prediction. Since Beaver's work in 1966, different prediction techniques with different variables, but mainly financial ratios, are applied both in takeover and failure predictions. As in the case of any other event in finance, information is power and enables its holders to manipulate the market and benefit from it. Thus, being in possession of early information about a takeover event is a major advantage to the holder.

As earlier event studies show, takeovers create abnormal returns to the target firm shareholders (Mandelker, 1974; Dodd, 1980; Conn, 1985; Holland and Hodgkinson, 1994). As studies analysed share price performance of companies around the announcement of bids, results were implicit that target firms' shareholders experienced abnormal returns (cumulative average residual) on their share prices after it was adjusted for overall market effect. However, the results were not so conclusive for bidding firms' share price performance. Results vary from positive returns to zero returns.

Jensen and Ruback (1983) examined 40 previous merger studies and concluded that corporate takeovers generate positive gains to target firm's shareholders, and bidding firm's shareholders do not lose. These results are, however, contradicted by some other researchers in that, although target firm's shareholders benefit in the form of premiums paid, acquiring firm's shareholders do not gain from the takeover activities. Roll (1994), by simply examining the same studies as Jensen and Ruback (1983), argued that the gains from takeovers are overestimated and reached a different conclusion, stating that the combined value of target and bidding firms has increased in some studies and decreased in others and none of these are statistically significant. Roll (1994) further argued that gains observed by the acquired firms represent a transfer from the bidding firms in the form of takeover premiums paid. Roll (1994) argues that corporate takeovers are an area that does not reflect the aggregate rational behaviour of markets but irrational decision making of individuals under uncertainty. As a result he explains that takeovers occur as a result of valuation error of individuals and/or hubris.

Some of the other studies reported that, on average, acquiring firms suffered falls in their share prices on the announcement of the takeover (Firth, 1980) while some indicated slight increases or no change in the acquiring firms' share prices. Asquith, Bruner and Mullins (1983) stated the importance of size effects and merger programs. They drew some attention to merger programs contending that earlier bids should contain more information about the profitability of the program than later bids. Their conclusion for the programs was that bidding firms gain significantly in their first four merger bids. Also, it is difficult to measure abnormal returns if the relative size of two merging firms is different. Nevertheless the overall conclusion of these studies is that takeovers are generally value enhancing.

Since it is generally supported empirically that target firm's shareholders benefit from a takeover it is especially important for market players to attempt to predict, if possible, or identify likely takeover targets in advance with certain degree of accuracy to acquire some abnormal returns.

1.2 The Purpose of the Study

This study undertakes an empirical analysis of takeover prediction in the UK. The main objective of the research is to examine whether it is possible to predict takeovers, or identify possible target candidates, before the actual announcement of takeover bid. As the question rightly put forward by Singh A. (1971), *"for firms quoted on the stock market, is it possible to generalise and suggest that the possession of certain definite economic and financial characteristics may make a firm more likely to be taken over?"* Second, if such a prediction is likely then more importantly to improve the prediction success of earlier studies by improving firm specific variables and employing a different technique that has started becoming an established analytical tool by it is extensive application in corporate finance. However it should be noted that it is not in the scope of this research to explain why takeovers occur, or to explain extensively the factors contributing to takeovers. As explained, the limit will be to examine whether company takeovers can be predicted by using some their publicly available information.

As mentioned, identifying target companies accurately prior to their bid announcement can create abnormal returns to the market players. Generation of abnormal returns to the target companies through corporate takeovers has already been well documented in the literature. However, contrary to markets' inability to make such predictions through valuation, Dodd and Ruback (1977), Asquith (1983), those studies mentioned above have already reported some impressive results by applying conventional statistical techniques of Multiple Discriminant Analysis (MDA) and Logistic Regression (LR).

Since the main aim of the study is prediction, a new technique, Artificial Neural Networks (ANNs), which has been tested and applied in other fields of finance and accounting, is put to use. The close parallel to the employment of ANNs to corporate takeover prediction studies in the field of corporate finance is corporate failure predictions. Several studies successfully use this technique in corporate failure prediction along with the classical counterparts of MDA and LR and reported that application of ANNs provided better classification rates than the former techniques (Alici, 1996).

Therefore this research will be carried out to test the following objectives:

1. Modelling takeover likelihood and testing it on the possibility of predicting target companies in the UK by ANNs and LR.
2. Applying ANNs in takeover prediction and testing ANNs' classification ability for target prediction in comparison with LR

1.3 Classification Techniques in Corporate Takeovers

It is worth noting that although, in this study, LR and ANNs are selected for classification techniques, classification techniques are by no means limited to these methods. Other techniques such as classification trees, multiple regression, probit analysis, machine learning techniques and expert systems can also be applied to the classification tasks (Weiss and Kulikowski, 1991). On theother hand, the reason for the preference given to LR, as a benchmark to ANNs, is its theoretical advantage concerning that it is free of multivariate normality assumption as well as its popularity in the literature. Also using the same type of technique will make the results of this study comparable to previous classification and prediction studies.

ANNs, as compared to the other statistical techniques, is relatively a new technique and making its way into financial applications. They are being applied to all areas of finance and accounting. Their rising popularity comes from the fact that ANNs try to mimic the human brain in their working. Thus, the approach of neural computing is to capture the governing principles of the human brain and its solutions and apply these methodologies to the given tasks.

Artificial Neural Networks (ANNs) have been used as a tool to classify companies or financial organisations (Martin-del-Brio and Serrano-Cinca, 1995) and much work has been conducted on bankruptcy predictions with the comparison of MDA and LR which are linear or curvi-linear classification techniques.

Alici (1996) has applied ANNs to company failure prediction and suggested that ANNs can perform better than their classical counterparts, DA and LR, where a noisy and random environment exists.

Several advantages of ANNs have been mentioned over statistical models. For example, this approach is free of multivariate normality assumption. In as much as the multivariate normality of financial ratios is questionable, the reliability of the models that use financial ratios and depend on this assumption of normality will be questionable too (Watson, 1990). The existence of outliers in financial ratios is widespread. Although remedial measures, such as winsoration, the replacement of outliers with the nearest observation, trimming, removal of unusual observations from the data, are applied to financial ratios, the distribution of many financial ratios is not only normal but their distribution properties change in time. However, it is claimed that ANNs are capable of fitting complex non-linear models to the data.

Recently, Sen *et al.* (1995) have applied ANNs and Logistic Regression to corporate merger prediction. Their work suggests that ANNs provide a good mathematical fit to the data. However in predicting mergers neither of the techniques performed better on a hold-out sample.

One of the main disadvantages of ANNs lies in their inability to explain the relative importance of the inputs. Therefore they have a limited use compared with statistical models in empirical research in finance and accounting. In order to eliminate this disadvantage, Sen *et al.*(1995) used simple techniques, such as sensitivity analysis and graphical plots. Also, Alici(1996) used a skeletonisation algorithm that aims to determine the relevance of individual nodes and connections by leaving only those important inputs and connections in order to eliminate this disadvantage to a certain degree.

Some of the applications of ANNs in finance and accounting are;

i. Bond rating: Dutta and Shekar (1988), Singleton (1990), Dutta and Wong (1994).
ii. Bank failure prediction: Bell *et al.*, (1990), Tam (1991), Tam and Kiang (1992).
iii. Stock price performance modelling: Kimoto *et al.*, (1990), Ahmadi (1990), Yoon and Swales (1991), Refenes *et al.*, (1994).
iv. Risk assessment of mortgage applications: Colins *et al.*, (1988), Reilly *et al.*, (1991).
v. Currency exchange rate forecasting: Refenes (1993).
vi. Time series prediction of Financial Markets: Sharda and Patil (1993), Azoff (1994), Beastaens *et al.*, (1994), Bosarge (1991), Beastaens *et al.*, (1994), Refenes (1995).
vii. Commodity trading: Bergerson and Wunsch (1991), Collard (1993).
viii. Accounting and financial ratio applications: Liang *et al.*, (1992), Coakley and Brown (1993), Berry and Triquerios (1993), Barker (1993), Kryzanowski *et al.*, (1995).
ix. Corporate failure prediction: Odom and Sharda (1990), Tam (1991), Udo (1993), Altman *et al.*, (1994), Wilson and Sharda (1994), Alici (1996).
x. Corporate takeover prediction: Sen *et al.*, (1995).

1.4 Data Set

The trend of mergers and acquisitions during the time period where the data set has been extracted can be seen below. This trend shows that the takeover activity reached a high point during the mid to late 1980s in the UK. It reached its peak point in 1987 in terms of numbers then in value terms in 1989. However, the trend has since been a decline. Even though it seems that it started to be on the increase in money terms, reaching a value of £36 billion in 1995, it declined to £23 billion in 1996. In the period of this study, from 1990 to 1994, the general trend is downwards both in value and number terms. This trend, especially in the drop in number of acquisitions, is obviously reflected in the data set of this study as well.

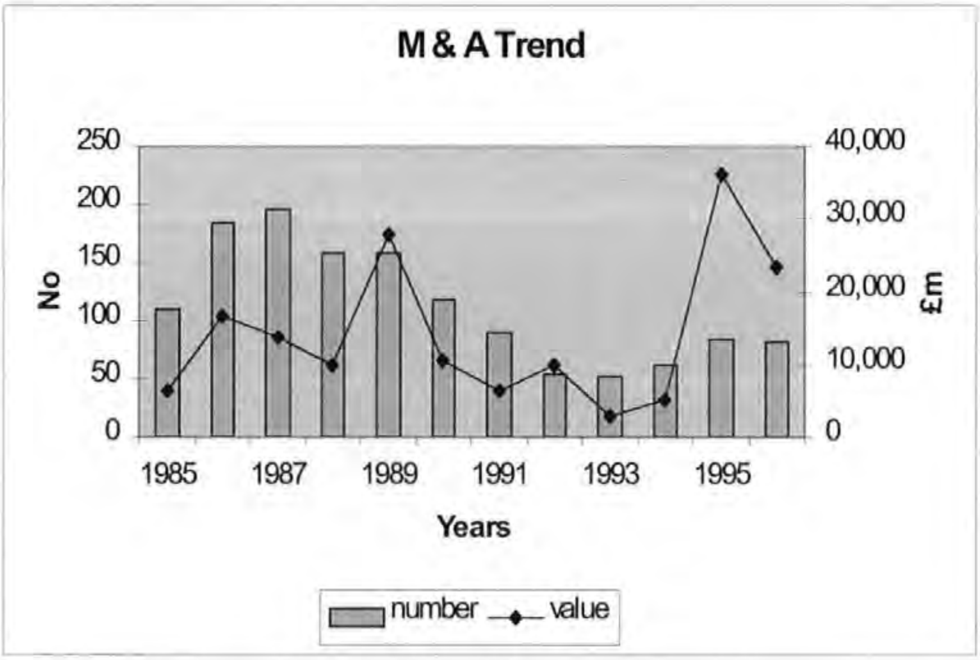


Figure 1.1. Number and Value of UK public company acquisitions.

The target companies have been extracted from the Acquisition Monthly from a period of January 1990 to December 1996 in the UK. The target group also includes failed bids. The assumption here is that if a company receives a bid, it contains the characteristics of target companies. The companies in the sample were drawn from manufacturing industry to establish more stable models by excluding oil and gas, utilities and construction industries. The target group includes 103 companies. The data set has been divided into two to form a derivation sample to form the models and a hold-out sample to test the effectiveness of these models.

In this study no further attempt has been made to split the target sample into hostile and non-hostile acquisitions. The fundamental reason for this is that the hostile acquisitions form a small proportion of overall takeover market in the UK. Weir (1997), for example, reported that out of 71 targets for the period 1990-93, only 23 were hostile acquisitions. This makes a proportion of 24% and is consistent with the figure for the 1980s (Sudarsanam, 1991; Powell, 1997). Small number of hostile targets would result models with unreliable estimates. Especially considering that the data is further divided into estimation and hold-out samples. If the same proportionality were applied to the target data in this study, it would have produced a sample of 24 hostile acquisitions. These 24 hostile companies would have additionally had to be partitioned into estimation and hold-out samples. Hence such a sample would have neither produced liable models nor prediction results would have been reliable. Therefore it is decided to concentrate solely on the whole acquisitions without splitting the targets into any sub-groups.

Two different sets of data are constructed in order to form the models. First the target group is divided into two random groups for each year in order to form the first model. 52 target companies are matched with non-target companies and used in the estimation sample. The rest of the target group is used in the hold-out sample. This data is called 'Mixed Data' (MD). Second, the target companies that received takeover bids from January 1990 to December 1994 are matched with *non*-target companies and used in the estimation sample. The target companies of 1995 and 1996 on the other hand are included in the hold-out sample. This data is called 'Time Data' (TD).

The reason for forming MD modelling is twofold. First it is used as an estimation of acquisition probability in the previous literature (Dietrich and Sorensen, 1984). Second to use it as a benchmark to TD. As can be seen the TD is a more realistic approach in takeover modelling. As it would be unrealistic to predict companies acquired in 1990 with companies takenover in 1996 and claim predictive success as a result in the case of MD.

The distribution of target companies in years for MD and TD are presented below in Tables 1.1 and 1.2.

Table 1.1-Target composition in MD

Years	1990	1991	1992	1993	1994	1995	1996	Total
No of Comp.	16	18	15	13	16	15	10	103
Derivation	9	9	7	6	8	8	5	52
Hold-Out	7	9	8	7	8	7	5	52

Table 1.2-Target composition in TD

Years	1990	1991	1992	1993	1994	1995	1996	Total
No of Comp.	16	18	15	13	16	15	10	103
Derivation	16	18	15	13	16	-	-	78
Hold-Out	-	-	-	-	-	15	10	25

1.5 Variable Set

Two different variable sets are used in this study. In the first part the financial ratios were used to model takeover likelihood. Financial ratios of the companies are collected from Financial Analysis Made Easy (FAME) and Datastream databases. Secondly the non-financial characteristics, which are the corporate governance structure, of companies are used in the modelling process. The corporate governance variables are collected from Price Waterhouse Corporate Register and annual reports of the related companies. In the final analysis the financial and non-financial variables are combined to form the models. The schematic description of this methodology is shown below.

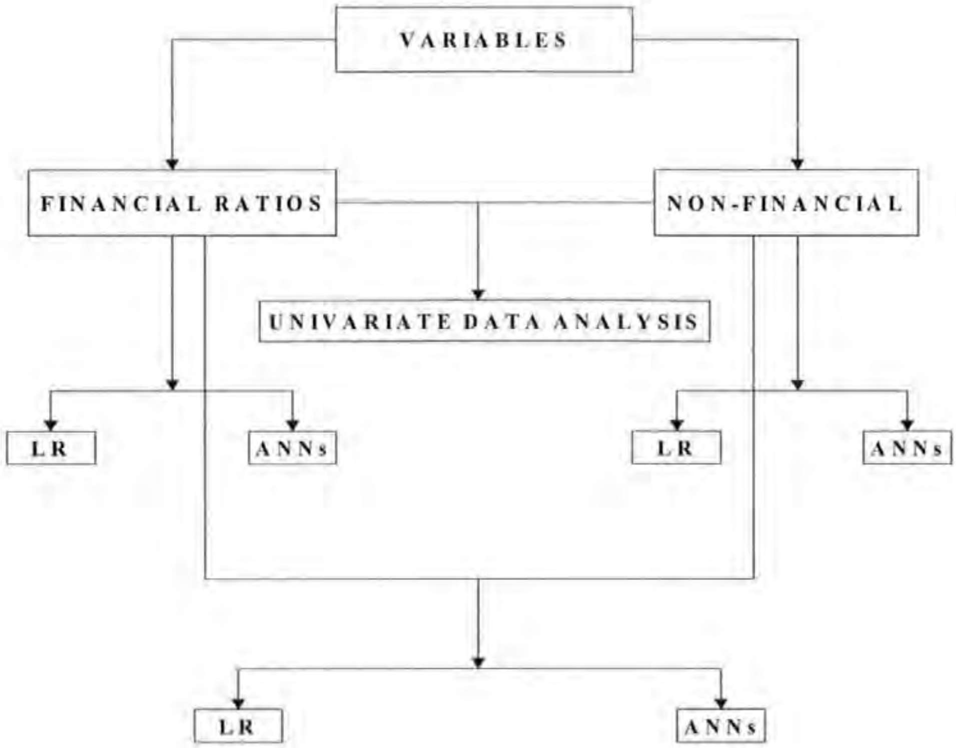


Diagram1.1 Schematic description of the steps involved in the methodology.

1.6 The Layout of the Study

In chapter 2, some theoretical discussion of mergers and acquisitions will be provided. In this chapter, the discussion on takeovers will be examined firstly from the two opposing views as neo-classical and managerial and in second from the agency theory perspective. The general purpose of this chapter is just to highlight the ongoing discussion on takeovers and takeover framework and provide an introduction to the concepts and terminology that are used in chapters four and five.

In chapter 3, the techniques that are used in this study will be covered. LR and especially ANNs will be described in more detail. Since one of the research objectives of this study is to measure the application of ANNs in takeover predictions, and due to ANNs' relatively recent introduction to finance, a relatively detailed introduction to neural computing technology will be given. Supervised learning process and training algorithms of Standard Backpropagation and Generalised Delta Rule along with unsupervised learning process are described.

In chapter 4, the financial ratio modelling results will be presented. The aim of this chapter is to model financial ratios of the companies in the sample and apply these models to hold-out samples in order to assess the prediction success. In the first section the methodology, data and derivation of the variables from the relevant hypothesis postulated in the literature as the causes of takeovers will be explained along with the methodological issues and shortcomings of prediction studies as explained by Zmijewski (1984) and Palepu (1986). In the second part the estimated models of LR and ANNs will be presented. Since there is no attempt will be made to exphin the relative importance of input variables in the constructed ANN structures, the classification tables of the networks' on the estimation sample will be given as an indication of the networks' mathematical fit to the data. In any case, the determination of the network weights is beyond the capability of the neural network software (NEUframe Professional v.3) that is used in this study. The final section of this chapter will provide the prediction results of the models and their significance.

In chapter 5, the corporate governance data modelling results will be given. Similar to chapter four, the first section describes the methodology, data and corporate governance variables. In the second part the estimated models of LR and ANNs' classification tables on the estimation data will be displayed. Finally, the prediction results of corporate governance models and their significance will be provided.

In chapter 6, the combined data modelling and results will be presented. As it is displayed in Diagram 1.1 that in the final stage of the analysis the financial and non-financial variables will be combined and modelled for prediction purposes.

In chapter 7, the main conclusions of the analysis will be discussed along with the limitations of the study as well as the general characteristics of takeovers that cause impediments to takeover predictions. Also the areas of further research to improve the ANNs modelling in takeover prediction studies or in similar empirical studies are discussed.

The reason for modelling financial ratio and corporate governance data separately initially before combining them in chapter six is to follow the paths of two distinct line of takeover likelihood modelling in the related research literature. Financial ratios used to estimate takeover likelihood extensively by the researchers such as Simkowitz and Monroe, 1971; Stevens, 1973; Singh, 1975; Belkaoui, 1978; Wansley and Lane, 1983; Dietrich and Sorensen, 1984; Rege, 1984; Walkling, 1985; Palepu, 1986; Barnes, 1990; Sen *et al,* 1995; Powell, 1997; and Barnes, 1998. On the other hand some other researchers such as Shivdasani (1993), Gammie and Gammie (1996) and Weir (1997) applied corporate governance characteristics of companies in their modelling. Therefore it is intended in this study to measure the predictive powers of these two separate firm characteristics independently without any interaction between them affecting the overall results.

1.7. Discussions and Summary

In addition to LR, ANNs will be applied to both financial and non-financial characteristics of takenover UK public companies in the hope that their unique profiles can be captured or modelled. These models will then be applied for discrimination and prediction purposes to test their fitness for such a task.

As far as the study is concerned, this will be the first study in which ANNs will be applied to takeover predictions by employing financial ratios in the UK and first time by employing corporate governance variables in takeover predictions.

It will also be valuable to see in the end how ANNs cope with such a complex task of takeover prediction and perform against an established multivariate technique such as LR.

Chapter 2 – Motives for Takeovers, Agency Theory and Control Environment

2.1 Introduction

The aim of this chapter is to provide a brief discussion of some of the theories of firm and their implications as well as their explanations of mergers and acquisitions process. In the first instance, the motives and reasons for mergers will be highlighted from the perspective of two opposing theories, classified as neo-classical and managerial approaches to firm behaviour. In addition, the results of some of the empirical work carried out to see whether acquisitions are undertaken in accordance with neo-classical or managerial perspectives will be reviewed. Furthermore, along with the discussion of these hypotheses, some of the variables and their relevance to the discussion topic will be highlighted. However, a more detailed discussion of the variables, financial and corporate governance, and their importance in terms of the takeover prediction perspective will be provided later in their respective chapters.

In the following part of the chapter, agency theory will be reviewed and its relevance to takeovers will be stressed. The review of agency theory in this chapter serves simply to highlight the agency framework, define the concepts and terminology that are used in the theory as well as in this study. It is not the aim here to provide a comprehensive or detailed analysis of agency theory in the sense of designing an agency concept. This review, though simple, provides a supplementary introduction and groundwork to the hypothesis that is used in the later chapters of the model development.

2.2 Motives and Reasons for Mergers and Acquisitions

Mergers and acquisitions decisions are part of business and corporate strategy for companies. However, the decision taken by the management to implement these strategies has brought forward the two different conflicting views on the motives of the takeover decisions. First it is the value maximising approach, where the bidding management makes a decision on the basis of the economic return to the investment, hence maximising shareholders' value, and the second is the managerial self-interest seeking approach.

In the neo-classical approach to the firm, managerial decisions about takeovers are taken to maximise shareholders' wealth. The neo-classical theory interprets the companies as single units with a single purpose of profit maximisation. It assumes costless information and rational decision-making so it isolates organisations from managerial behaviour. Accordingly self-interest seeking purpose has no place. It further assumes that power relations are unimportant and the conflict among the individuals is meaningless (Moschandreas, 1994). The neo-classical profit maximisation theory suggests that competitive market forces drive companies to pursue shareholder wealth maximisation. Hence, the theory interprets mergers and acquisitions from a profit maximisation perspective and concludes that companies make acquisitions when it results in increased shareholder wealth. If, after the takeover, the acquired company's profitability increases, then this increase in shareholder wealth will be realised (Firth, 1980; Manne, 1965). Firth (1980) suggests that the profitability can increase through monopoly power, synergy, or replacing target management with a more efficient and competent type.

Jensen and Ruback's (1983) review of the literature is supportive of the neo-classical theory. The results of the reviewed empirical studies showed that shareholders of target firms accumulated large positive abnormal returns in completed takeovers while small but statistically significant positive abnormal returns are realised by bidders in successful tender offers. However the returns to bidding companies' shareholders in the case of mergers were zero. Finally, Jensen and Ruback (1983) concluded that even though bidding companies' shareholders do not gain but do not lose either, gains experienced by the targets indicate that takeovers create value.

Mandelker (1974) investigates mergers and acquisitions from a competition perspective to see whether the information about the mergers is efficiently reflected in the company's share price. The study concluded that shareholders of acquiring firms earn normal returns from acquisitions as they would from other investment activities with the same risk level. It also reported that acquired firms earn abnormal returns after the merger. The results that acquired firms' shareholders are not losing but gaining positive normal returns support the profit maximisation hypothesis.

Langetieg (1978), examining mergers from the same perspective as Mandelker (1974), reached the same conclusions. It is reported that acquiring firms' shareholders obtained non-significant but positive returns from mergers and acquired firms had an average excess return of approximately 13%. These results are also in support of the profit maximisation purpose of mergers. However, the research also concluded that since gains to the acquiring firms are too small, it is possible that the profit maximisation was not the only reason for mergers.

However, the results of various other researches showing that target shareholders are benefiting, and acquirer shareholders are either losing or not gaining, implicated and challenged the neo-classical perspective of acquisitions. They have concluded that managers pursue takeovers in order to promote their own self-interests and some of these self-interest-seeking factors are listed as, a reduction in the risk of losing a job increased salary levels and power, or diversification and minimisation of bankruptcy risk. Mueller (1969), for example, argued that managerial salaries, bonuses, stock options, and promotions tend to be related more to changes in the size of a firm than its profits.

The managerial perspective looks at the issue from the separation of management from ownership perspective and infers that the managers as individuals act to enhance their own interest rather than that of their shareholders. The theory suggests that managers are motivated by salaries and other financial rewards, and desire to increase their status and power. Berle and Means (1932) were the first to highlight the separation of ownership and management and its consequence for the profit maximisation purpose. They have perceived that management and ownership interest do not naturally align when the roles are separated. This however may not necessarily mean that the profit is totally sacrificed; it is also suggested in the theory that after a certain satisfactory profit level, the management will attempt to maximise their own utility. Otherwise they would have been dismissed by the shareholders.

From a point of view of takeovers, the self-interest seeking behaviour of managers will result in unprofitable acquisitions to increase size and have a downward effect on share prices and, as a result, on shareholder wealth unless proper mechanismsare put in place to curb and control the self-interest seeking actions of managers. These mechanisms will be discussed in further detail in the second part of this chapter.

Firth's (1980) study of the period 1969-75 concluded that even if the motive of mergers were to create profit maximisation and shareholder wealth, the acquiring companies' shareholders experienced wealth losses. Firth (1980) also measured the relation between remuneration levels and increased size after a merger activity and found that while takeovers have resulted in losses to shareholders, it benefited the directors in monetary terms, hence supporting the self-interest seeking motivation theory of takeovers. Firth (1991) moreover found that after the acquisition the managers' pay increases alongside the increase of the acquirer's size. This increase is substantial when the share value of the bidder increases but also when it decreases. Similar results were also observed by Meeks and Whittington (1975) on a UK based sample. They found that sales growth is positively related to pay increases.

Jensen and Murphy's (1990) research on performance pay and top-management incentives concludes that the relation between CEO wealth and shareholder wealth has been in decline and is small. This finding suggests that the managers do not have incentives to act in the best interest of their shareholders.

Dodd (1980) studied the daily market reaction to announcement and subsequent acceptance or rejection of merger proposals. It is reported that for completed merger proposals, target firms' shareholders earn positive abnormal returns in contrast to cancelled proposals where they earn significant negative abnormal returns. However shareholders of bidding firms in either case earn negative abnormal returns from the merger proposals. These findings provide support for a managerial perspective of acquisitions.

Strong empirical support for the managerial perspective came from Agrawal *et al.*, (1992). In their study, they have looked at long term performance of acquiring firms after mergers and even after adjusting for the size effect and beta risk. They have found that shareholders of acquiring firms sustain a statistically significant loss over the five year term after the merger date. Their results also show that negative returns experienced by acquiring firms are not as a result of a slow adjustment of the market to the takeover event.

Most recent UK evidence on the managerial wealth maximisation hypothesis came from Gregory (1997). He finds that after controlling for size and book-to-market effects, acquiring firms, even though they had a significant out-performance three years previous to the merger activity, were significant under-performers three years after the merger date. He also reported that post acquisition losses were more severe in cases where the acquisitions were financed by equity rather than by cash.

One other reason suggested by Sudarsanam (1995) for the failure of acquisitions creating value for the acquiring shareholders was that mergers and takeovers are subject to different intra-organisational conflicts and political decisions, which are likely to result in poor acquisition decisions.

From an investment point of view, mergers and acquisitions are investment decisions of a kind that provide an external growth as opposed to an internal one and are evaluated in the same way as any other investment decision. They are preferred to organic growth decisions at times when there is a need for rapid market expansion or defending existing markets, speedy diversification into new markets to exploit existing or potential profit opportunities through target's assets and market share (Hutchinson, 1995; Sudarsanam, 1995). Copeland and Weston (1988) argue that growth is vital to the well being of companies, as it is needed to attract the best managerial talent by offering rapid promotions and responsibilities and without a constant flow of skilled managers, the companies would decline in value. Mergers also provide a means of rationalising and consolidating in declining industries (Chiplin and Wright, 1987).

As already mentioned, the causes for mergers and acquisitions cover wide, diverse and complex reasons. These could be synergy, risk reduction, strategic reasons, business cycles, the economic and political environment, integration into European Union, self-interest seeking management motives, diversification, defence against takeovers, and financial motivations such as taxation and increased leverage.

These wide and diverse reasons for mergers and acquisitions can further be clustered into two broad categories for the purpose of this study. These can be grouped broadly as industry specific or related factors (including synergy effects and the impact of restructuring) and corporate control and agency cost motivations (Dickerson *et al.*, 1998).

Gort (1969) highlighted the significance of industry specific factors as important motives for acquisitions. He observed that merger frequencies vary among the industries and distribution of acquired and acquiring companies is not uniform among industries but concentrates on certain types of industries. He, following this trend, discarded the hypothesis that mergers occur as a result of managers' personal ambitions and increase their asset base for security reasons or simply to manage bigger firms or any external influences such as tax structures.

One of the main industry specific factors is synergy. Synergy has been cited as one of the main reasons for takeovers and for value creation in most of the merger activities. Synergy can be achieved through reductions in production or distribution costs as it creates economies of scale in operational. The acclaimed benefit of a merger where there is synergy is as commonly stated as '2+2=5'. In an efficient market, the merged organisation is valued in such a way that the added value gained through the merger is added to the sum of the individual parts. This states that synergy creates a value to the combined entity so that the whole is bigger than the sum of separate parts. In other words, the value of the merged firm exceeds the value of the individual parts. Sudarsanam *et al.*, (1996) reported that synergy creates value for shareholders of targets and bidders where there is complementary growth-resource imbalance between bidders and targets. Synergy reason or claim is especially obvious when the intended merger or takeover is horizontal or vertical. Vertical or horizontal integration, for example, is used by companies in order to achieve economies of scale or to reduce competition in a given industry, prevent newcomers, or drive benefits from a dominant or monopoly power.

A second factor in the industry specific group is the economic shocks (Gort, 1969), where through technology or rapid changes in the security prices, that increase the merger activity by increasing the dispersion in valuations. Following an economic or technological shock, for example, mergers and acquisitions could be used as a medium to restructure in specific industries. In the event of a technological shock in a given industry, capacity reduction might be required and could be achieved through acquisitions. Mitchell and Mulherin (1996) analysed industry level patterns in takeover and restructuring activity in 1980s over 51 industries and reported that those industries experienced the greatest amount of takeover activity during the study period are the ones that exposed to the greatest fundamental shocks. These fundamental shocks are defined as deregulation, changes in input costs, and innovations in financial technology that induce or enable alterations in industry structure. Jensen (1994) also argues that as a result of changing technology or market conditions takeovers occur in order to structure corporate assets. In order to control such effects researchers used industry dummies, and measured size and liquidity effects. Palepu (1986) and Sen *et al.,* (1995) applied industry dummy variable in their analysis. They assigned a value of one if at least one acquisition observed in a firm's four digit SIC industry during the year prior to the year of observation. On the other hand Powell (1997) did not consider industry dummy in his analysis of takeover modelling in the UK. Industry dummies will not be used in this study since there is at least one merger activity in all the four digit industry classifications in the time period analysed.

Sudarsanam *et al.,*(1996) show for a UK sample of takeovers that synergy - operational, managerial and financial -- creates value for shareholders of targets, bidders and in some cases, where there is complementary liquidity slack and surplus, for both of them. They have also found some evidence that ownership structure of bidders and targets has a significant effect on their shareholders' wealth. Conclusions in their study very much depended on the synergy effect and, in conditions where highly rated companies acquired less highly rated companies, acquiring companies' shareholders incurred wealth losses whereas target shareholders experienced wealth gains.

If the proposed merger is not in a common/same industry then it is a conglomerate merger. These types of mergers were quite common during the 60s as part of risk reduction strategy by the companies. The diversified entity, through conglomerate merger, might lead to a reduced risk for the combined entity and provide easier access to financial capital and markets. Mueller (1969) argued that conglomerate mergers might be explained by the existence of management synergies. However, his argument has been challenged, as in a perfect functioning capital market such risk reduction can quite easily be achieved by shareholders on their own by simply holding a portfolio of shares. It is also claimed that conglomerate mergers are less likely to be successful as the acquiring management is not familiar with the targets' market. Agrawal et al., (1992) found that acquirers of conglomerate mergers had negative performance over the five year post merger period. However contrary to expectations, the loss to the non-conglomerate acquirers was higher.

In contrast to Gort's (1969) argument of dismissing managerial ambitions to increase the size of their companies, the study by Amihud and Lev (1981) showed that, in the case of conglomerate mergers, managers engage in takeover activities in order to reduce their employment risk as they cannot diversify it. However, there are some studies to challenge their conclusions otherwise. Matsusaka (1993), by examining the stock market response to the conglomerate merger wave of the 1960s, found that acquirer's shareholders benefited from these diversification mergers and as a result concluded that these merger activities were not driven by managerial motives.

In contrast to industry specific factors, the market for control hypothesis provides some different explanations for the reasons for mergers and acquisitions. The main point of the argument of the market for control hypothesis is poor performance. Managers who pursue their own interest rather than shareholders' will be replaced by other competing management teams. Fama and Jensen (1983) defined corporate control as *'the right to determine the management of corporate resources – that is, the right to hire, fire and set the compensation of top managers'.* A market for corporate control is defined as a market where different managerial teams compete for the rights to manage corporate resources (Jensen and Ruback, 1983).

In theory the stock market will assign a low share price to companies with low profitability and low expectations of future performance. This will reduce the possibility of obtaining new financing from the market. The main indicator of such poor performance is profitability or how efficiently the assets of a company have been utilised to generate high levels of profitability. Manne (1965) argued that the share price of a company is the main indicator of a management's performance. In order to measure the poor performance of a management or poor profitability, in this study, return on shareholders fund (RSHF) will be employed. So it is expected that if the managers are pursuing their own benefits instead of that of their shareholders then this will be reflected in the profitability and as a result in their share price. Also utilised is market-to-book value (MBV) and price earning ratio (P/E). As MBV and P/E are estimated by using share price information, they can be used to assess, in part, the markets' disciplinary involvement in acquisitions. Therefore a market for corporate control will discipline the actions of non-value-maximising managers through takeovers. Hence lower profitability should increase the probability density of such companies for possible takeovers. Dickerson *et al.,* (1998) suggested two other metrics for the measurement of how corporate control identifies those non-value optimising managers. These are dividend payments and investments.

Dividend payments and investment policy metrics are the natural extension to the free cash flow hypothesis postulated by Jensen (1986). Free cash flow is the cash available to a company after it carries out, after meeting its tax commitments, all available positive net present value (NPV) investments. Jensen (1986) explains that 'free cash flow is cash flow in excess of that required to fund all projects that have positive net present values when discounted at the relevant cost of capital'. The hypothesis implies that the companies that lack any profitable investment opportunities should distribute their free cash to their shareholders as dividends or use it to purchase their own shares. Hence the free cash flow will not be used for value reducing investment schemes. However, those who use it for further investment will be disciplined by the market, as their likelihood for takeovers will increase. However the results from the Dickerson *et al.*,(1998) study found no evidence that firms without profitable investment opportunities are more likely to be taken over in the event that they increase investment or reduce dividends.

2.3 Agency Theory

In modern corporations, the separation of management from ownership has highlighted the conflict between agents and principals. Agency theory argues that under conditions of incomplete information and uncertainty, which are the main characteristics of the business environment, two agency problems arise: adverse selection and moral hazard. Adverse selection is the condition under which the principal cannot ascertain ex ante if the agent accurately represents his ability to do the work for which he is being paid. Moral hazard is the condition under which the principal cannot be sure if the agent has put forth maximal effort (Eisenhardt, 1989). The conflict of interests between management and shareholders and description of this relationship through contractual agreements between the two parties are at the centre of agency theory. Simply, agency theory is concerned with the principal-agent problem in the separation of ownership and control of a firm (Jensen and Meckling, 1976; Hill and Jones, 1992), and in particular the problem rooted of aligning the interests of two groups through contractual agreements in order to maximise the firm performance.

In principle, the agents ought to act in the best interest of their shareholders and try to maximise their returns. In reality, though managers do not have the same opportunity to diversify their financial wealth in the same way as their shareholders do, they have invested their human and financial capital in one company that their career reputations depend on its performance. As a result they might end up acting not necessarily in the best interest of their shareholders but their own.

Overall, the agency theory attempts to resolve different objectives and risk preferences of principals and agents, within an agency relationship. In the first instant the problem is not only the different objectives and goals of principals and agents but also the difficulty of verifying the actual aims of agents by principals (asymmetric information). A second problem arises due to the different risk preferences of the two parties (Eisenhardt, 1989). As put forward by Gerald and Stout, (1992), the nexus of contracts is the centre of agency theory and it focuses on the inherent conflict between the interest of shareholders and the interest of those who run and work for the firm.

An agency relationship is defined by Jensen and Meckling (1976) as '*a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision making authority to the agent*'. In general, as explained above, agency theory defines a relationship between two groups who have different goals and attitudes toward risk (Eisenhardt, 1989).

These interests of managers as managerial motives are almost equally and similarly recognised by different theories of the firm as financial rewards, status, prestige, power, and security (Williamson, 1964). This is consistent with self-interest seeking motivation and the existence of asymmetric information that complicates monitoring by shareholders. The occurrence of this conflict of interest between principal and agent was raised by Adam Smith as early as 1776:

> *"The directors of such companies...being the managers rather of other people's money than of their own, it cannot well be expected, that they should watch over*

it with the same vigilance with which the partners in a private copartnery frequently watch over their own. Like the stewards of a rich man, they are apt to consider attention to small matters as not of their master's honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company."

The process of diversification of shareholding and diffusion of management ownership in modem corporations has created wide share dispersion so that no one individual or small group of shareholders holds a sufficiently substantial amount to exert significant influence on the day to day running of the businesses. This has left managements wih insignificant shareholdings in the businesses that they run. Berle and Means (1932) concluded as early as 1932 that after examining 200 large US corporations, wide share dispersion might divert the interest of owners and managers from each other. Accordng to their research result no shareholder had as much as 1 per cent of total shares and that the twenty largest shareholders had no more than 5.1 per cent of total shares. On the other hand, the shareholdings of institutional investors has been on the rise over the years. This trend of share dispersion and increasing shareholdings of insurance companies can also be observed in the UK (see Table 2.1).

Table 2.1-Ownership of UK equities as percentage, 1969-95

Sector of Beneficial owner	1969	1975	1981	1991	1992	1994
Individuals	47.4	37.5	28.2	19.9	20.4	20.3
Other personal sector (mainly charities)	2.1	2.3	2.2	2.4	1.8	1.3
Public sector	2.6	3.6	3.0	1.3	1.8	0.8
Ind. & commercial companies	5.4	3.0	5.1	3.3	1.8	1.1
Oversees	6.6	5.6	3.6	12.8	13.1	16.3
Banks	1.7	0.7	0.3	0.2	0.5	0.4
Insurance comp.	12.2	15.9	20.5	20.8	19.5	21.9
Pension funds	9.0	16.8	26.7	31.3	32.4	27.8
Unit trusts	2.9	4.1	3.6	5.7	6.2	6.8
Other financial institutions (mainly investment trusts)	10.1	10.5	6.8	2.3	2.5	3.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Central Statistical Office, *Share Ownership* (1995) 8.

It is also argued that due to this wide diversion of shareholding, a small percentage of shareholders may effectively influence or control the affairs of a company through voting power. If these small but influential groups of shares are held by the managers then they can effectively pursue their own interest rather than that of the shareholders. However, recent growth and dominance of institutional investors and their active involvement in firm-level monitoring through routine dialogue can be an active device to affect corporate behaviour and can also act as an incentive for managers to pursue the interest of shareholders (Moschandreas, 1994, Stapledon, 1996). This assumption, of course, holds true to the extent that institutional investors are willing to participate in the corporate affairs of their equity investments. They may see their participation as an extra burden on themselves and employ the easy option of selling their shares on the market in cases where they are not satisfied with the incumbent management rather than participating in the decision making process.

As for the perspective of takeovers, agency theory argues that the threat of a takeover provides a permanent mechanism to monitor management performance and discourage or reduce the pursuit of management interests at the expense of shareholders. In the event of failure of the internal control mechanisms, as boards may be dominated by individuals who owe their positions to their relations rather than merits, 'cronyism,' and thus the takeover market- as an objective, market-based mechanism - is an essential tool in a system of contracts (Jensen, 1993). Walkling and Long (1984) provide some insights into this conflict by analysing cash tender offers and management resistance to takeover. They concluded, for example, that management resistance to takeover attempts is significantly affected by the wealth chance of officers and directors.

The dominant rationale for takeovers in agency theory is poor performance of targets. In the financial economic approach to the firm, the firm's share price provides the only objective indicator of management performance (Manne, 1965). More efficient management is recognised by the market and reflected in share prices. On the other hand, low share price will provide the means for other management teams to acquire the resources/assets of these companies. This process not only provides an opportunity for outsiders who detect undervalued firms and an economic safety net for shareholders; moreover, it provides a mechanism to discipline top managers that fail to serve shareholder interests, as they usually find that their employment terminated following a takeover.

Besides the market for corporate control, there exist certain natural, regulatory and contractually designed mechanisms on the market to align the interest of shareholders with managers and reduce the self-interest seeking behaviour of agents. It is worthwhile to examine some of these mechanisms which have been introduced or already existed in order to minimise the managers' incentives to deviate from their contractual purpose of profit maximisation and eventually shareholder wealth maximisation. Since some of these mechanisms will be used as a justification to include some of the variables that are put to use in the following chapters, it is worthwhile examining them here.

2.4 Agency Cost and Control Environment

Agency theory demonstrates that divergence from the shareholder wealth maximisation creates an agency cost. As it is mentioned earlier, in an agency-principal relationship, an agency problem arises due to the divergence of interest between the two parties. In an ideal world where the actions of agents or outside factors influencing an outcome can be observed perfectly and information is available without any incurring cost, agents act in the best interest of principals, and any deviations from the ideal create agency problems and agency costs. However, due to the reasons outlined above, in reality managers have incentives to deviate from their contractual obligations. Managers not only pursue their benefits through perks such as company cars, expensive office items, etc., but through selecting investment, operating, or financial policies that are suitable to their personal risk and time preferences (Byrd *et al.*, 1998). The conflicts arising from a principal-agency relationship are called agency problems and the loss of value due to these problems is called an agency cost.

It can be summarised here that an agency cost arises from (i) a divergence of interest between principal and agent which in short can be defined as a self-interest seeking behaviour of both parties (ii) an existence of asymmetric information which complicates the monitoring process of the agent's actions. Agency cost comprises of (i) monitoring costs incurred by principals to monitor agents to make sure that they direct their efforts in the principal's interest, (ii) bonding costs incurred by agents to bond themselves to act in the interest of principals, and (iii) residual lost incur because not all actions of agents can be monitored.

Four types of possible causes of these problems are identified. These are (Byrd *et al.*, 1998):

1. Effort: managers might spend less time with business than might be expected from them by shareholders.
2. Horizon: managers' horizons for expected returns tend to be shorter than for shareholders.
3. Differential risk preference: unlike shareholders, managers cannot diversify their wealth and have a different attitude to risk.
4. Asset use: managers have incentives to misuse a company's assets or consume excessive perks.

Hence, decisions and objectives taken by the management would not necessarily be aiming to maximise shareholders' value and it has been argued that necessary steps should be taken to minimise this divergence by aligning the interest of agents with principals. Jensen and Meckling (1976), for example, suggested that by establishing appropriate incentives for the agent and by incurring monitoring costs designed to limit the divergent activities of agents, the problem of divergent interests of two parties can be overcome. Nonetheless, even with these incentives and costs there may still be some actions of agents that will differ from the actions that principals would take. This is defined as residual cost. On the other hand, according to Fama (1980), agency costs are driven to zero by market forces both within and external to the firm. A form of full ex post settling up occurs which penalises self-interest seeking managers and causes them to act in the best interest of shareholders.

Fama and Jensen (1983a) provide an extensive insight into an organisation's decision-making process as well as a framework to reduce agency problems. An organisation's decision process consists of decision management (initiation and implementation) and decision control (ratification and monitoring) (Fama and Jensen 1983b). Separation of residual risk bearers (owners) and the decision functions (management) leads to an important agency problem. In order to prevent or minimise this problem, initiation and implementation functions of decision making should be separated from the ratification and controlling functions.

As suggested by Fama and Jensen (1983a), devices for separating decision management from decision control include:

1. Decision hierarchies, in which the decision initiatives of lower agents are passed on to higher level agents for ratification and monitoring.
2. Boards of directors that ratify and monitor the organisation's most important decisions hire, fire and compensate top level decision managers.
3. Incentive structures that encourage mutual monitoring among decision agents.

As outlined by Stapledon (1996), many monitoring mechanisms exist and are applied by shareholders to reduce the conflict of interest. In addition to these available mechanisms, the external monitoring environment is also in place to exert pressure on management. The final goal of these internal monitoring devices, such as a Performance Related Pay Scheme, is to cope with agency problems. Some of these mechanisms are:

(i) Market Forces: The most important of all market forces is the market for corporate control by takeovers. The market for corporate control theory was first put forward by Manne (1965), providing the theoretical foundations for most of the research in takeovers thereafter. According to Manne (1965), a fundamental premise underlying the market for corporate control is the existence of a high relationship between corporate managerial efficiency and the market price of shares of that company. Therefore the stock market provides an objective evaluation of management performance through the price it places on a firm's equity. As a result, inefficient management would be evaluated on the market by the equity price that will eventually create an incentive for others to take control of the firm for better utilisation of the firm's assets. Obviously, a high share price of an equity is not only to deter other firms from bidding due to the cost, but can also be seen as a sign that the market is quite happy with the current management and its way of running the firm's assets. Takeovers, as discussed above, redistribute the corporate assets from poorly performing teams to more efficient ones. Agency theory argues that the existence of a takeover market as a threat is an extra incentive for management teams not to divert from the profit maximisation objective. Furthermore, capital markets exert pressure on incumbent management by making it difficult for them to raise equity capital or debt financing. The higher pricing of securities of the successful firms enables them to have cheaper access to investment funds and the lower pricing for the lesser successful ones will have a penalising effect. On the other hand, however, this idealistic pricing of securities depends on accuracy of forecasting of future profits and there is a good deal of evidence that the prices yielded by stock markets are quite far from such perfection. Share prices are known to be influenced by asymmetric information, speculation and other market imperfections. Third, the product market plays a monitoring role as the company may lose its market share as a result of inefficient management and may become insolvent. Finally, the market for managerial talent puts pressure on the management team to run the company successfully so it can open up its job prospects within the market.

In their study of "Market for Corporate Control", Jensen and Ruback (1983) discussed corporate control and viewed it as an arena where management teams compete for corporate resources for their better utilisation. Therefore, through this competition those management teams who cannot utilise their company resources efficiently will be eliminated by the other competing teams in the market in order to maximise resources.

(ii)Executive Incentive Remuneration/Performance Related Pay Scheme: this theory suggests that, linking executives' remuneration to the share performance or to changes in shareholder wealth provides the means for aligning the interests of the two groups and reduces the managerial self interest seeking behaviour.

(iii) Mandatory Disclosure of Information and Auditing: the requirement to disclose periodic financial and non-financial information and auditing of this information by independent auditing firms provide the means for the shareholders to assess the condition of the company and discourages management diverting from the shareholders' profit maximisation.

(iv) Non-executive Directors: it is accepted that non-executive directors on the board can monitor management as outside observers. Their presence on the board may deter management from diverting from their primary responsibilities.

(v) Shareholder/Large block of Shareholder Monitoring: shareholders of companies, as laid down by the Companies Act 1985, have the right to remove directors without a reason. Therefore, any management team that diverts from the profit maximisation goal may eventually be fired by the unsatisfied shareholders. Large shareholders have an incentive to monitor managers as they benefit more from monitoring. Unlike small shareholders they are less likely to be free riders.

There are also other mechanisms such as regulatory investigations and inspections of Directors' and officers' duties. Regulators have statutory powers by Company Directors Disqualification Act 1986 to investigate the affairs of companies and take actions such as disqualifying directors.

2.5 Discussions and Summary

Although acquisition literature has come up with some theoretical explanations on motives and reasons for mergers and acquisitions, the empirical evidence on these hypotheses has not produced a conclusive result to accept or dismiss one or the other explanation. It has been reviewed above that empirical research is deeply divided on the motives of mergers.

Research shows that target shareholders benefit from acquisitions by the premiums paid irrespective of the acquirers' motives. It is obvious that acquirers' motives are to be the determinant factor as to whether their shareholders benefit from the acquisition activity. Sudarsanam *et al.* (1996), in their research, for example, clearly demonstrate that when an acquiring companies' management engages in takeovers in order to utilise synergies, their shareholders gain as well as the target shareholders. On the other hand when their actions are clouded with hubris or motivated by self-interest seeking then an acquisition can have a negative effect on share price and there will be a reduction of shareholders' value.

Diffusion of ownership has benefits as well as drawbacks to the modern corporations. While the day to day running of large corporations has been improved in the hands of professional managers, their lack of involvement in some aspects of risk sharing, short and long term expectations, and small or non-existent share interest have created a conflict of interest with the shareholders. This has led to agency problems and cost.

In the face of these problems, several incentive and control mechanisms that operate either independently or in conjunction with each other are designed to reduce or at best eliminate agency costs. The characteristics of these mechanisms vary with the existing risk and uncertainty. However, even with application of these internal and existence of external devices, it is hard to assume that the managers' self interest-seeking behaviour is totally eliminated and the agency cost is driven to zero.

As outlined above, the failure or lack of application of these mechanisms or the failure in product market or the market for managers will further decrease management efficiency if they pursue their own interest and reduce firm value which will be reflected in the share price. In a market place where management teams compete for better utilisation of corporate resources, the market for corporate control, which is often referred to as the takeover market, will act and shift these resources for better utilisation.

However, it is also likely that in a market where stock prices show random behaviour, some of the companies may become undervalued and become a target. This occurs not because managers have failed to maximise profits or as a result of failure in any of the control mechanisms or markets, but through the stock market's random pricing. This will create opportunities for those that are overvalued through this process and give them the incentive to make a bid.

As mentioned earlier, the monitoring or control environment is important from a takeover perspective since they form the basis of some of the assumptions for efficient design of a company management structure on a corporate governance level and their implementation to minimise the agency costs.

In the event that some of these monitoring mechanisms are not in place, or are inefficiently designed, this may lead to agency costs and a lower value of a firm. The absence or internal design and implementation of these mechanisms within the company structures will be analysed and exploited from a takeover prediction perspective in this study.

Chapter 3 – Classification Methods in Takeover Prediction Studies

3.1 Introduction

Theoretical background to mergers and acquisitions and it relevance to this study has been explained in the previous chapter. The aim of this chapter is to provide similar background information to the statistical techniques that were used in the takeover likelihood modelling as well as to present some derivation and application details of the techniques applied in this study.

The main statistical techniques that have been used in the previous empirical studies, which were carried out to predict those companies which are a target or likely to receive take-over bids, are Discriminant Analysis (DA), Logistic Regression (LR) and recently Artificial Neural Networks (ANNs). Only in one of these studies that Sen*et al.*, (1995), applied ANNs to takeover predictions. Some of these studies can be listed as, Simkowitz and Monroe (1971), Stevens (1973), Singh (1975), Belkoui (1978), Wansley and Lane (1983), Dietrich and Sorensen (1984), Palepu (1986), Bartley and Boardman (1990) and Barnes (1990), Sen *et al* (1995), Powell (1997).

Even though only LR and ANNs are applied in this study a brief description of DA along with LR and ANNs will be also provided in this section. Moreover the reasons for not selecting DA in this study as well as a comparison with LR will be explained. However, the emphasis will be on the explanation of ANNs, compared with LR and DA, as this is the emerging technique. This overview of concepts of ANNs will enhance the understanding of terms and structures of the technique in the following sections.

Furthermore in section 3.3 examples of research studies which used and applied ANNs in finance and accounting will be listed and their main results will be reported in addition to the list provided in chapter 1. The aim here is to see the type of applications of ANNs in finance and accounting as well to have an understanding of these techniques measured outcome against the traditional multivariate statistical models.

3.2 Classification Techniques

The literature on classification techniques is vast and these techniques can be divided as numeric and non-numeric (quantitative). Numeric techniques cover deterministic and statistical measures, while non-numeric techniques fall into the area of symbolic processing that is dealt with by such methods as fuzzy sets (Beale and Jackson, 1990).

Although the most commonly used classification techniques in takeover studies seem to be DA, and LR, there are numerous statistical techniques that are also available for classification purposes such as multiple regression, K nearest neighbour, classification trees, multidimensional scaling, machine learning and expert systems. Kennedy (1992), for example, has applied and compared the results of seven classification techniques, such as linear discriminant analysis, quadratic discriminant analysis, McKelvey and Zavoina n-chotomos probit, Walker and Duncan ordinal logit, Nerlove and Press polytomous logit, ordered classification trees, and unordered classification trees, that have been used in accounting research. Furthermore, Mar-Molinero and Ezzamel (1991) have applied multidimensional scaling technique to corporate failures. As can be seen, the classification techniques are by no means restricted to DA or LR.

The reason why LR, rather than the other ones mentioned above, is used in this study is that, as explained above, its widespread use in the field and many applications of ANNs are ultimately benchmarked against it in terms of performance (Tam and Kiang, 1992; Yoon *et al.*, 1993; Curram and Mingers, 1994; Serrano-Cinca, 1997; Miranda and Burgess, 1997; Luther, 1996; de Carvalho *et al.*, 1998).

3.2.1 Linear Discriminant Analysis (LDA)

Discriminant analysis is used to examine the differences between two or more naturally occurring groups of objects with respect to the related predictors. The basic assumption in LDA is that two or more naturally occurring groups exist and differ on several of their characteristics. (Klecka, 1980). In corporate takeover predictions, LDA is used to determine the group membership of companies as targets and non-targets from their given characteristics or discriminating variables. The employed predictors are generally the financial ratios of companies. Discriminating variables must be measured in interval or ratio level in order to satisfy the assumption of multivariate normal distribution of variables.

The most widely used DA method, Linear Discriminant Analysis (LDA), is the one developed by Fisher (1936). Linear discriminant analysis is the two-group case of Multiple Discriminant Analysis (MDA). The Fisher procedure builds a discriminant function by maximising the ratio of between groups and within group variances. Provided the following conditions are satisfied the Fisher procedure minimises the expected cost of misclassification.

The assumptions are;

1. Non of the variable should be a linear combination of another variable. Hence the correlated variables should be excluded from the analysis.
2. Population covariance matrices are equal for each group.
3. Each group is extracted from a population which has a multivariate normal distribution. Multivariate normal distribution holds that the scores on predictors (discriminators) are randomly distributed, and that sampling distribution is of any linear combination of predictors and is linearly distributed. In the event this assumption is violated, the estimated probabilities are not exact.

The discriminant function is a linear combination of independent variables. The function assigns a score, Z, to each company in a sample through the discriminating characteristics of the companies. Discriminant function has the following mathematical form.

$$Z = b_0 + b_1X_1 + \ldots\ldots\ldots + b_nX_n \qquad (3.1)$$

Where,

b_l = discriminant coefficients,

X_l = independent discriminating variables,

Z = the score on the discriminant function for observation.

In order to allocate new observations to previously defined groups, a cut-off score is established such that companies can be assigned to groups with respect to their scores. The ones that have scores below the cut-off will be considered as possible targets and vice versa. In the estimation of weights the score (Z) variance between two groups (target and bidder) is maximised relative to the variance in scores within groups.

Simply DA classifies companies into one of the groups (target or non-target) on the basis of individual scores (Z). These scores are the weighted combination of identified characteristics that best separates targets and non-targets.

As mentioned above this technique assumes that independent variables are multivariate normal and covariance matrices of two groups are equal. Nevertheless, the assumption about the multivariate normal distribution has frequently been ignored and violated in the applications. The assumption is especially violated when a dummy independent variable is used in the analysis. On the other hand, it is suggested that some remedial actions can reduce the severity of this violation. Altman, Haldeman and Narayanan (1977) improved the normality of asset size variable by using logarithmic transformation. Even though various researchers claimed that the DA is relatively robust to non-normality, Eisenbeis (1977) stated that violation of normality assumption may bias the test of significance and estimated error rates.

The second assumption states that the covariance matrices across the variables are equal for target and non-target companies. Eisenbeis (1977) explained that the significance test for the differences in group means is effected when this assumption is ignored or relaxed. It is suggested (Tam and Kiang, 1992) that if the covariance matrices are different, quadratic functions can be used rather than linear functions. However it is pointed out that even though quadratic classifiers are quite accurate in classifying the training sample, they perform (their generalisation ability) poorly on hold-out samples compared with linear models (Altman *et al.*, 1981; Sudarsanam and Taffler, 1985).

Even though, as stated generally by researchers, the discriminant analysis is relatively robust to non-normality, it is highly sensitive to outliers.

DA has been quite popular as a statistical technique in finance and accounting and applied for predictive purposes by employing accounting data. One of the most popular and related areas to takeover predictions where DA is applied is bankruptcy prediction. In takeover studies DA is used to estimate a linear model that best discriminates between targets and non-targets. The model is then applied to determine group membership of companies as targets and non-targets.

In order to eliminate the above assumptions about the population, for example other forms of DA can be utilised. As suggested by Curram and Mingers (1994) kernel DA or nearest neighbour DA do not make assumptions about the underlying populations. However they have also stated that these two procedures are essentially heuristic procedures and do not provide consistent performance.

As a linear classifier LDA is trying to find a straight line that separates two groups. This is causing some shortcomings for the technique. One of the common problems that can be faced with the application of LDA can be exemplified in "exclusive-or" problem.

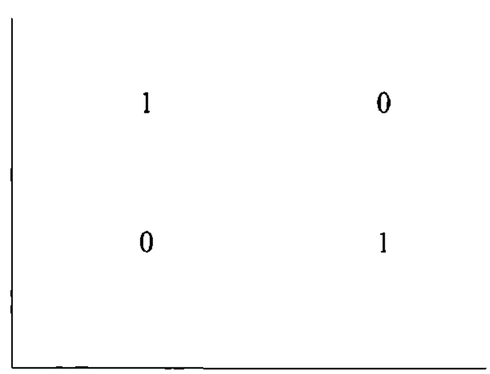
The "exclusive-or" problem is summarised below (Yoon *et al.*, 1993; Beale and Jackson, 1990).

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

Where X_1 and X_2 are input variables. Y is the output.

In this case the output is 0 if the inputs are (0,0) or (1,1). While the inputs are (1,0) or (0,1) the output is 1. When a linear function ($Z = b_0 + b_1X_1 + b_2X_2$) is formed to model above example with a cut-off point Z^*, the "exclusive-or" problem is unavoidable. If X = (0,0) and X = (0,1) then $b_2 > 0$. Similarly $b_1 > 0$ if X = (0,0) and X = (0,1). If both b_1 and b_2 are positive, when X = (1,1) then $Z^* < b_0 + b_1 + b_2$ as a result X = (1,1) will not be classified as group 0.

Graphically the "exclusive-or" problem can be depicted as;



As can be seen the pattern is linearly inseparable since no straight line can be drawn to separate two groups. This situation can arise in circumstances where non-linear relationship exists. Simple example of this can be drawn out from Palepu's 1986 paper where growth-resource imbalance variable would fit into above graph. If horizontal axis represents growth, such as turnover growth, while vertical axis represents resource, such as liquidity, then 1 on the graph represents companies that have growth-resource imbalance problem. In the mean time zeros indicated those companies that do not have the imbalance problem. These two groups of companies can not be separated by a straight line.

3.2.2 Logistic Regression (LR)

LR has been another multivariate statistical model widely used in empirical research. It is an extension of multiple regression analysis techniques and follows the same general principles as linear regression. The difference between the two models is the outcome variable. Unlike in the linear regression where the dependent variable is continuous, the outcome variable is binary or dichotomous in logistic regression. This difference between the two modelling approaches is reflected both in the choice of a parametric model and in the assumptions (Hosmer and Lemeshow, 1989).

Although LR performs the same task as LDA, there are some differences between these two techniques. LR uses a sigmoid function that provides an output between 0 and 1 and very suitable for studies on takeover or bankruptcy predictions. Also LR estimates the coefficients through a probabilistic method based on maximum likelihood while DA mostly uses Wilks's Lambda to estimate them. This approach of probabilistic estimation of coefficients means that LR is free of the assumptions of DA about the underlying assumptions of population.

The logistic model is formulated as;

$$\Pr(Y=1) = \frac{1}{1+\exp\left[-\left(B_0 + \sum_{i=1}^{k} B_i X_i\right)\right]} \tag{3.2}$$

where the response variable takes only one of two possibilities. The mathematical expression on the right side the logistic model formulae is the genaral logistic function.

$$f(z) = \frac{1}{1+e^{-z}}$$

Where $z = B_0 + \sum_{i=1}^{k} B_i X_i$. Here the $f(z)$ function ranges from 0 to 1 as z various from -∞ to +∞. Graphical depiction of the function is also shown in figure 3.1.

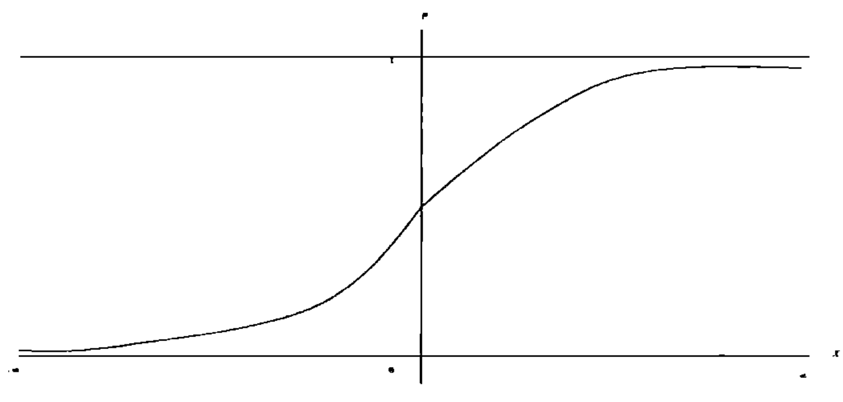


Figure 3.1. Logistic function.

LR provides the conditional probability of an observation belonging to a certain class, given the values of independent variables (covariates) for the observation. It is based on the cumulative probability function and does not require the multivariate normality of the covariates. The general method of estimating the model parameters is called maximum likelihood (Maddala, 1989; Hosmer and Lemeshow, 1989; Kleinbaum *et al.* 1998).

Maximum likelihood estimation refers to a general algorithm for obtaining estimators of population parameters. When the underlying distribution of the dependent variable is normal then maximum likelihood estimation is identical to least squared estimation. However when compared to least squares method maximum likelihood can be applied to linear as well as nonlinear models. Maximum likelihood estimation can be explained with an example of the problem at hand, simply the estimation of covariates of takeover likelihood. This discussion below is based on Kleinbaum *et al.* (1998) explanation of maximum likelihood.

If the population contains a certain unknown proportion $\theta(0 \leq \theta \leq 1)$ of some companies with characteristics of receiving a bid and a random sample of m companies is sampled from the population, there will be random number of companies, Y, with the characteristics of receiving a bid in the random sample of m. The possible values of Y are the $m+1$ integer values of $0,1,2,\ldots,m$. The estimation problem here is to use Y and m to obtain an estimate of θ.

Above the underlying probability distribution of the discrete random variable Y is binomial and for an observed value of Y, the maximum likelihood estimator of θ is the value of $\hat{\theta}$, for which the expression (3.3) attains its maximum value as a function of θ.

$$pr(Y;\hat{\theta}) = {}_mC_Y \theta^Y (1-\theta)^{m-Y} \qquad (3.3)$$

where ${}_mC_Y = m!/Y!(m-Y)!$.

The specific value of $\hat{\theta}$ that maximises (3.3) can be obtained by taking the derivative of (3.3) with respect to θ. The resulting equation gives,

$$\frac{d}{d\theta}\left[pr(Y;\theta)\right] = {}_mC_Y \theta^{Y-1}(1-\theta)^{m-Y-1}(Y-m\theta) \qquad (3.4)$$

Equating (3.4) to 0 provides three solutions 0,1 and Y/m. Here 0 and 1 minimise equation (3.4) while Y/m maximises. The second derivative of (3.4) with respect to θ is negative when it is replaced by the value of Y/m. Thus the maximum likelihood estimator of θ is $\hat{\theta} = Y/m$ when Y has the binomial distribution. Here $\hat{\theta}$ is the sample proportion of companies with characteristics of receiving a bid.

The maximum likelihood estimator $\hat{\theta} = Y/m$ has the property that pr(Y; $\hat{\theta}$) > pr(Y; θ^*). Where θ^* is any other value of θ satisfying $0 \le \theta^* \le 1$. As pr(Y; θ) gives the probability of observed Y, the estimator $\hat{\theta} = Y/m$ is called the maximum likelihood estimator of θ and most agrees/fits with the data.

In a general sense the method of maximum likelihood gives values for the unknown parameters that maximises the probability of obtaining the observed data. However in order to apply this method first the likelihood function (function of the unknown parameters) is constructed. This function represents the likelihood of observing the data or gives the probability of the observed data as a function of the unknown parameters. Once the likelihood function is determined, the maximum likelihood estimates that estimator of the set of unknown parameters which maximises the likelihood function. Thus, the resulting estimators are those which most agree/fit with the data (Hosmer and Lemeshow, 1989; Kleinbaum, 1994).

The cumulative probability assumption is needed in order to constrain the predicted values to be in the range of zero-one (0=non-target, 1=takenover) (Figure 3.1). The coefficient of each variable can be interpreted as the effect of unit change in an independent variable on the probability of the dependent variable (Zavgren, 1983). Thus, the parameters of logistic regression give the importance and sensitivity of each predictor in the model.

In our modelling of takeovers, logistic regression weights the independent variables and creates a score (y) for each company. The observed score then used to calculate the probability of group membership.

$$y = B_0 + B_1X_{i1} + B_2X_{i2} + \cdots + B_nX_{in} \tag{3.5}$$

$$\text{Probability of M \& A } (P(y = 1)) = \frac{\exp(B_0 + B_1X_{i1} + B_2X_{i2} + \ldots)}{1 + \exp(B_0 + B_1X_{i1} + B_2X_{i2} + \ldots)} \tag{3.6}$$

Where $X_{i1},..........,X_{in}$ are predictor variables. As explained above, the B_i are estimated by maximising the likelihood function (3.7). More detailed derivation of likelihood function can be seen in Kleinbaum *et al.* (1998).

$$L(Y;B)=\frac{\prod_{i=1}^{n1}\exp\left(B_0+\sum_{j=1}^{k}B_jX_{ij}\right)}{\prod_{i=1}^{n}\left[1+\exp\left(B_0+\sum_{j=1}^{k}B_jX_{ij}\right)\right]} \tag{3.7}$$

Where the B is the vector of parameters to be estimated. The maximisation leads to a series of nonlinear equations that are solved for B using iterative mechanisms. Once B is estimated group probabilities are estimated using (3.5) for a given set of variables, X_i.

The main advantage of applying LR, for example compared with Multiple Discriminant Analysis, is that, as mentioned above, it does not depend on the assumptions of multivariate normality and equal covariance matrices for each group. Given the violation of these assumptions in accounting and finance data, application of LR as a benchmark against ANNs seems most appropriate rather then MDA (Eisenbeis, 1977). The imposed normality assumption on the independent variables is clearly violated when dichotomous or categorical variables are used in the model. As put forward by Kleinbaum (1994) *" if any of the independent variables are dichotomous or categorical in nature, then the discriminant function method tends to give biased results, usually giving estimated odds ratios that are too high"*. On the other hand, maximum likelihood estimation does not require any restriction of any kind on independent variables. Therefore the independent variables can be ordinal or nominal.

As suggested by Kleinbaum, (1994), one of the main reasons why DA has been applied extensively in financial modelling until recently even with the widespread violations of the fundamental assumptions of DA, was the absence of available computer software and power for the estimation of maximum likelihood.

3.2.3 Artificial Neural Networks (ANNs)

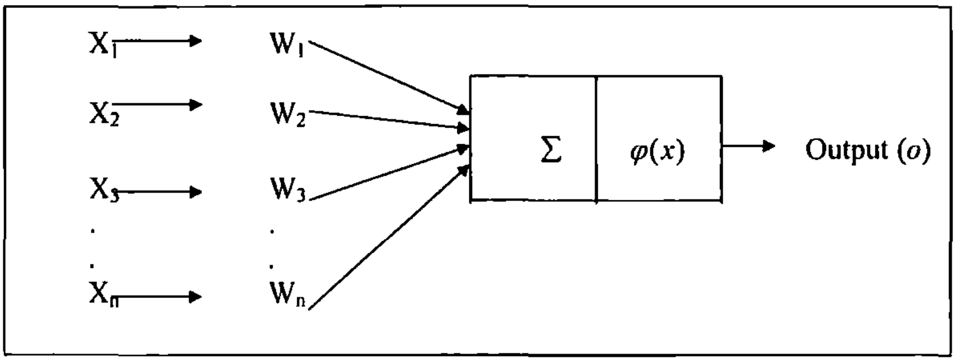
The concept of ANNs is developed from the workings of the human brain and its different computational abilities. Although ANNs are developed from the workings of the brain, it is not, however, the exact replica of it. Their aim is to mimic the brains network structure not to replicate it.

A neural network is made up of layers of information processing units called neurons. Each neuron performs a simple weighted sum of the information it receives. The weightings or coefficients are called synaptic weights in neural network terminology. An activation/transfer function is applied and an output obtained which, in turn, outputs as an input to another layer of neurons. The transfer function is usually sigmoid, but it can also be a linear function. The schematic description of this process can be seen below in Figure 3.2.

3.2.3.1 Artificial Neuron

A neuron is an information-processing unit of the ANNs and its basic function is to add its inputs, and if this summed inputs exceed a value known as threshold value, to produce an output value. The working of a neuron is depicted in Figure 3.2.

Figure 3.2 Artificial Neuron



As can be seen in Figure 3.2, an artificial neuron receives inputs as data and firstly sums them up and then squashes them by the activation/transfer function $\varphi(x)$. Then this value is compared with the threshold value to produce an outcome or output (O), if the sum of the inputs exceed the threshold value, otherwise it does not. This system of transforming and passing the inputs to produce an output is called feedforward system.

The working of a neuron can be formulated in mathematical terms taking the following form:

$$O = \sum_{i=1}^{n} XiWi$$

Where: O = linearly combined output

Xi = input values

Wi = weightings

n = number of observations.

This sum is then compared to the threshold value and if the sum is greater then the output is produced. It can also be formulated as:

$$O = \sum_{i=1}^{n} XiWi - \theta$$

Where: θ = threshold value

Therefore the decision is taken to produce an outcome on the basis of whether the O exceeds the value of θ. The decision is taken on the following basis before activating transfer function:

Transferred	$O > \theta$
Not Transferred	$O <= \theta$

The process explained above can be reorganised so that the same result would be obtained. Instead of subtracting the threshold value, an additional input can be connected to the neuron. By doing this the extra input of +1 is multiplied by a weight equal to minus the threshold value, $-\theta$, and added with the other inputs. This is called biasing the neuron. This process is shown in the figure 3.3.

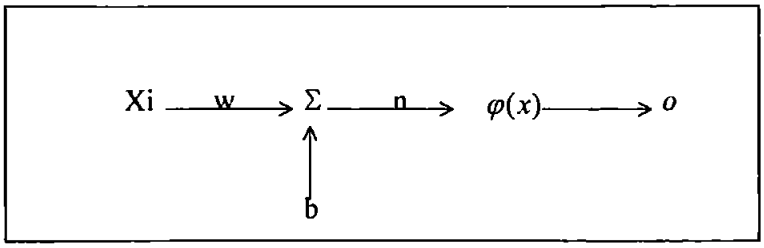


Figure 3.3 Artificial Neuron with Bias

Where: b = bias.

In this case the output is formulated as:

$$O = \varphi(\sum_{i=1}^{n} XiWi + bi)$$

or the formula can be rewritten as:

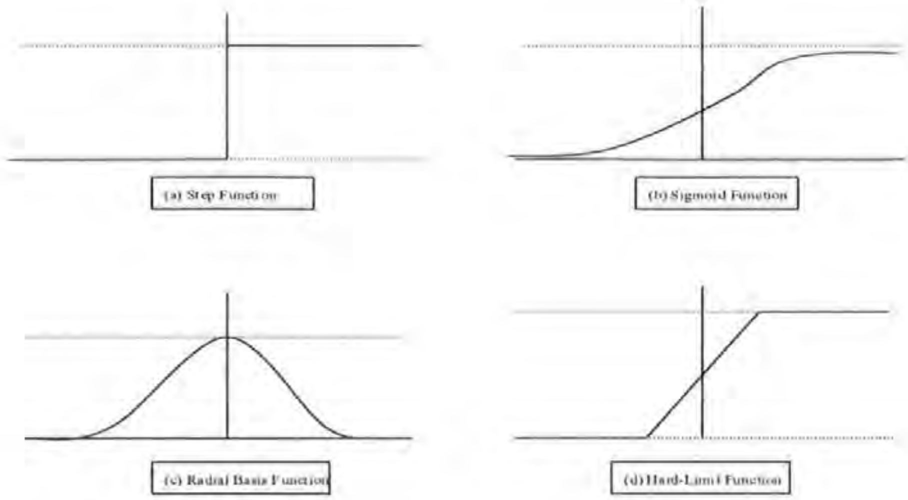
$$O = \varphi(\sum_{i=0}^{n} XiWi)$$

Where X_0 is set to zero.

3.2.3.2 Transfer Functions

Transformation to other neurons is determined by a non-linear function. In other words, the transformation function as denoted by $\varphi(x)$ determines how a neuron will scale its response to incoming data as it is considered before the transfer function produces an output. In addition to the step function below, transfer function can take some of the forms presented below in Figure 3.4 (adapted from Maren, 1990, p.48).

Figure 3.4 Neural Network Transfer Functions



Logistic sigmoid functions are often used as activation functions among the transfer functions (Bishop, 1995). Because of its s shape it represents a balance between linear and non-linear behaviour (Haykin, 1998). The reason for their popularity is that they are complex, differentiable, unlike the step function, and suitable for many applications, especially in finance (Refenes, 1995).

There are two types of sigmoid functions, symmetric and asymmetric and they take the following form:

$$O = 1 \ / \ 1 + \exp(-x)$$

3.2.3.3 The Perceptron and the Learning Rule.

A perceptron consists of an input and output layer without a hidden layer. Each neuron in the input layer is connected to each neuron in the output layer. Weights between the connections are adjusted as the perceptron tries to find an optimum solution to the problem presented.

Given any input pattern, a perceptron produces a set of output values from the neurons in the output layer, and the network's output is determined by the input pattern and the weights of the connections. At the beginning of the model estimation/training, the connections are initialised as small random values, so the output produced by the network is unlikely to exceed the output desired for any particular pattern. The connections are adjusted during the training so that the output produced by the network approaches the desired output as more information is fed into the system.

The basic principle is that the connections from the input neurons are unchanged if the output produced by the network is correct. However, if the output exceeds the desired output then the connections between the neurons are decreased, on the other hand the weights are increased if the output produced is smaller than the desired output.

The basic training, Perceptron Learning Algorithm (PLA), of the perceptron can be summarised as (Beale and Jackson, 1990):

> *Step 1*: Initialise weights and threshold (set the weights and thresholds randomly). Set W_i ($0 <= i <= n$) to small random variables where W_i is the

weight for input i, and θ to be the threshold value in the output node. Also, set W_0 to be $-\theta$, the bias, and b_i to be always 1.

Step 2: Present inputs and the desired outputs. Present inputs x_i ($0 <= i <= n$) with the desired outputs d.

Step 3: Calculate the actual output.

$$O = \varphi(\sum_{i=0}^{n} XiWi)$$

Step 4: Adapt weights to reduce the error to an acceptable level. Alter the weights so that the correct decisions are reinforced and the incorrect ones are corrected.

$$Wi = WiXi\eta(d - O)$$

Where η ($0 <= \eta <= 1$) is the positive gain term (learning rate) that controls the adaption rate.

This procedure is continuous until the error term (d - o) reaches a pre-set acceptable level. This learning algorithm is known as supervised learning as the correct answers are imposed on the net.

Another learning algorithm similar to this one is known as Delta Rule. In this case the adjustment of the weights is larger when the weighted sum deviates significantly from the desired value which is measured as $\Delta = (d - o)$, and the adjustment is smaller when the deviation is also smaller.

The perceptron, though able to discriminate between two linearly separable classes, is unable to solve non-linear problems. However this problem of the perceptron is overcome by the Multi-Layer Perceptrons (MLP). They differ from the perceptron in respect to number of hidden layers between the input layer and the output layer. The activation of hidden and output layer neurons is the same as in the case of a simple perceptron, and the response function is a non-linear function which is usually a sigmoid function.

3.2.3.4 Multi-Layer Perceptrons (MLP) and the Learning Rule.

Multi layer perceptrons are feed-forward nets with one or more layers of nodes between the input and output nodes (Lippmann, 1987). A simple MLP contains three layers: an input layer, a hidden layer and the output layer. Each layer in a MLP has a weight matrix, a bias vector and an output vector. This is presented in Figure 3.5.

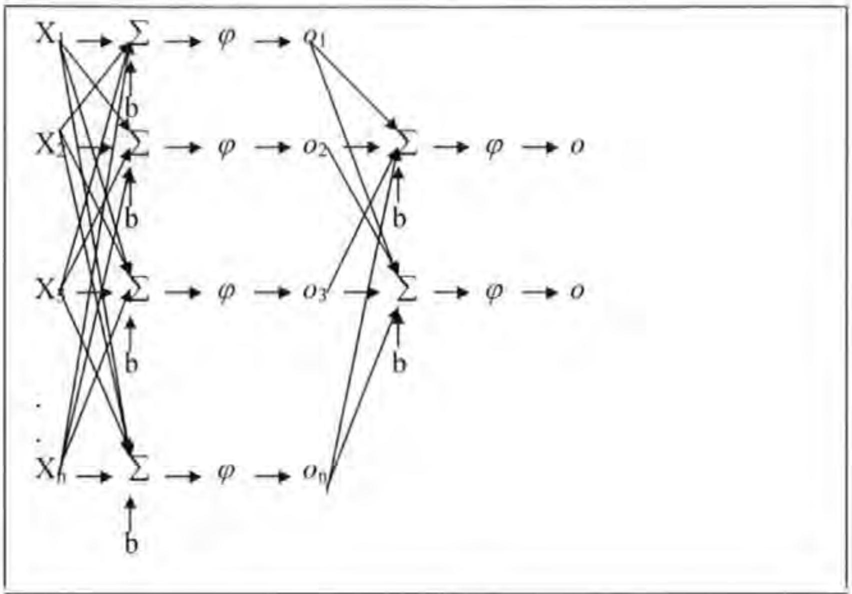


Figure 3.5 Multi-Layer Perceptrons

The network shown above has n inputs z number of neurons in the hidden layer and the two in the output layer. It is important to note that the number of neurons in the hidden layer do not have to be equal to the number of input units. MLP may contain more than one hidden layer and, if added, the neurons in the second layer may also be different in number than the first hidden layer. As can be seen outputs of each layer form the inputs of the following layer.

MLPs with just one hidden layer can learn to approximate any function to any degree of accuracy given enough hidden units, enough data and training time. As a result, MLPs are known to be universal approximators. Although one hidden layer is sufficient in most situations, there are cases where a network with two or more hidden layers may require fewer hidden units and weights than a network with one hidden layer, so using extra hidden layers sometimes helps to improve generalisation (SAS, 1999).

Training takes place in a similar manner as with the perceptron. As a start, the weights and biases are initialised to small random values. The training pattern is fed into the network from input units, and outputs of neurons in the first layer are calculated. The output, called activation, produced by these neurons is transmitted through the response function onto the neurons in the second layer. This process of forwarding is repeated until the final output is obtained.

As in the case of the perceptron, the output from the network will be subject to a learning algorithm. The weight connections will be adjusted until the desired output is achieved. This is achieved by the back pass during which the changes in the connections are back propagated (BP) through the network.

The BP algorithm is used to adapt the weights such that the final output of the net approaches the desired output. This process is achieved by determining the differences between the actual output and the desired outputs and using these differences to adjust the weights. However, compared to the weight adjustment in Perceptron Learning Algorithm (PLA), the desired output of the hidden units are unknown. Therefore, the desired and actual outputs of the output layer are used to change the weights by propagating it back through each hidden layer to minimise the error. It is suggested that that the hidden units that are connected to the outputs with large errors adjust their weights more than the ones with small errors. The mathematics shows that the weights for a particular node should be adjusted in direct proportion to the error in the units to which it is connected. This is the reason that back-propagating the errors in the net allows the weights to be adjusted correctly.

3.2.3.5 ANNs Learning Procedures.

ANNs use two types of learning procedures. One is the supervised learning and the other one is unsupervised learning known as Kohonen's self-organising feature maps. The two learning algorithms, PLA and BP, mentioned above are the examples of supervised learning where desired output is presented to the net so that the net learns from each input-output pair. Supervised learning is instruction-orientated. The objective of this type of learning is to minimise the cost function by eliminating the differences between the desired output and the actual output (Refenes, 1995). In supervised learning, the desired output is presented to the net so that the net learns from each input-output pair. An analogy can be made to a classroom situation where a teacher corrects the networks' response to its inputs. As inputs are fed in to the net, an output is produced and compared with the desired or correct output. The teacher forces the network to change its internal presentation of data in order to capture the essential features of the input data. The teacher determines the learning rule and the weight adjustments. This form of learning has been employed in the majority of neural network applications.

On the other hand, in an unsupervised learning procedure there is no such desired output presented to the network to minimise the cost function. Unsupervised learning is self-organisation orientated (Refenes, 1995). The objective of this learning rule is to model the features (cluster) in the data. Compared with the supervised learning in unsupervised learning, there is no teacher to correct the answer. Instead the network is left with the inputs and is expected to discover patterns in the training data by creating the most appropriate solution to the problem. This self-organising feature of neural networks involves competition or co-operation, or both, amongst the neurons.

3.2.3.6 ANN Algorithms

Based on the learning rules, there are different algorithms. The most common supervised learning algorithm, as already mentioned above and used in this study, is Standard Back Propagation Algorithm (SBP). Radial Basis Function (RBF) is another algorithm under the supervised learning. The reason for not applying RBF algorithm in this study is that it requires the training target to be continuous analogue data, not binary or categorical text (NCS, 1997).

The major algorithm that corresponds to unsupervised learning is the Kohonen Algorithm. Kohonen networks are one of the most widely used unsupervised learning techniques. In the case of unknown desired output for each training pattern, the network self-organises so as to group the training patterns into similar categories. As this algorithm is more suited to unknown desired outcome (clustering) it will not be used in this study.

A mathematical working of SBP Algorithm based on Lippmann (1987) and Beale and Jackson (1990) is discussed in the section below, as it is SBP algorithm that will be applied in this study. However, the more detailed mathematical explanations of the backpropagation algorithm can be seen in Rumelhart *et al.* (1992) and Beale and Jackson (1990).

3.2.3.6.1 Backpropagation and Generalised Delta Rule

The aim of the backpropagation algorithm is to reach convergence at a desired minimum level. To do this it uses the Generalised Delta Rule (GDR). In GDR, the convergence is achieved through a gradient decent procedure that uses incremental changes in the network's weights. Manipulating the learning rate and momentum controls the level of this change.

The general goal is to find the combinations of weightings which provide global minimum error. The first step is to calculate the minimum error for pattern $\boldsymbol{p}$, $\boldsymbol{E_p}$, in order to achieve target output for pattern $\boldsymbol{p}$, $\boldsymbol{t_{pj}}$. The next step is to compare this with the actual output for the same pattern, $\boldsymbol{o_{pj}}$.

$$Ep = \frac{1}{2}\sum_{j}(t_{pj} - o_{pj})^{2} \qquad (1)$$

This is a specific error function for pattern $\boldsymbol{p}$.

The activation for each node, $\boldsymbol{j}$, for pattern $\boldsymbol{p}$ becomes

$$net_{pj} = \sum_{i} w_{ij} o_{pj} \qquad (2)$$

This is equivalent to the weighted sum in the single layer perceptron.

As in the case of simple neuron, the weighted sum for each node, j, is passed onto the transfer function, f_j. The transfer function in MLP can be any of the above mentioned. As mentioned it is sigmoid activation function that is most commonly applied.

$$o_{pj} = f_j(net_{pj}) \tag{3}$$

The error function for pattern p, E_p, is determined by incremental changes in weights in the network. So, the derivative of error function for pattern p with respect to weightings, with chain rule, can be expressed as

$$\frac{\partial E_p}{\partial w_{ij}} = \frac{\partial E_p}{\partial net_{pj}} \frac{\partial net_{pj}}{\partial w_{ij}} \tag{4}$$

The first part shows the change in error with respect to the net input node whereas the second part measures the effect of the changes in the net with respect to particular weights on the net input.

Consider the second term in equation (4), and substituting in equation (2), the effect of the changes in the weights on the net can be expressed as

$$\frac{\partial net_{pj}}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \sum_k w_{kj} o_{pk} = \sum_k \frac{\partial w_{jk}}{\partial w_{ij}} o_{pk} = o_{pi} \tag{5}$$

since $\frac{\partial w_{kj}}{\partial w_{ij}} = 0$ except when k=i when it equals to 1.

Changes in error with respect to the change in the net inputs to a node can be defined as

$$-\frac{\partial E_p}{\partial net_{pj}} = \delta_{pj} \tag{6}$$

So equation (4) becomes

$$-\frac{\partial E_p}{\partial w_{ij}} = \delta_{pj} o_{pi} \tag{7}$$

In order to decrease the value of E_p, weight changes proportional to $\delta_{pj} o_{pi}$ are made. This is simply the implementation of the gradient descent in E, where η is the learning rate which determines the level of changes.

$$\Delta_p w_{ij} = \eta \delta_{pj} o_{pi} \tag{8}$$

it is δ_{pj} that should be estimated for each of the nodes to decrease the global error, E_p. Recalling equation (6), the chain rule can be written as

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}} = -\frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial net_{pj}} \tag{9}$$

The second term can take the following form as in equation (3)

$$\frac{\partial o_{pj}}{\partial net_{pj}} = f_j'(net_{pj}) \tag{10}$$

The first term in equation (9), recalling equation (1), by differentiating E_p, with respect to o_{pj}, becomes

$$\frac{\partial E_p}{\partial o_{pj}} = -(t_{pj} - o_{pj}) \tag{11}$$

Thus the error term for each node can be written as

$$\delta_{pj} = f_j'(net_{pj})(t_{pj} - o_{pj}) \tag{12}$$

As for the written nodes, the error term can be presented as follow, by using the chain rule,

$$\frac{\partial E_p}{\partial o_{pj}} = \sum_k \frac{\partial E_p}{\partial net_{pk}} \frac{\partial net_{pk}}{\partial o_{pj}}$$

$$= \sum_k \frac{\partial E_p}{\partial net_{pk}} \frac{\partial}{\partial o_{pk}} \sum_i w_{ik} o_{pi} \tag{13}$$

$$= -\sum_k \delta_{pk} w_{jk} \tag{14}$$

By using equation (2) and (6), and being aware that the sum decreases and the partial derivative is non-zero for only one value, such as in the case of equation (5). When a substitution between equation (14) and (9) takes place, the changes in the error function with respect to weights in the net can be written as

$$o_{pj} = f_j'(net_{pj})\sum_k \delta_{pk} w_{jk} \quad (15)$$

In summary, the equation (12) calculates the error in the output nodes and propagates back to the network to change the connection weights of previous nodes. This calculation and modification is done by equation (15). A multi-layer network uses equation (12) and (15) in parallel in a backpropagation learning framework.

3.2.3.6.2 Backpropagation Learning for Multi-Layer Perceptrons (MLPs)

The majority of MLPs uses sigmoid activation function that takes the form of

$$o_{pj} = \frac{1}{1 + e^{-(\sum_i w_{ji} o_{pi} + \theta_j)}} \quad (16)$$

where θ_j is a bias. In order to apply the learning rule, the derivative of the activation function with respect to its total input, ***net***$_{pj}$, where $net_{pj} = \sum w_{ji} o_{pi} + \theta_j$, should be known. When we take the derivative of O_{pj} with ***net***$_{pj}$,

$$\frac{\partial o_{pj}}{\partial net_{pj}} = o_{pj}(1 - o_{pj}) \quad (17)$$

For a logistic activation function, the error signal, δ_{pj}, for an output unit given, is

$$\delta_{pj} = (t_{pj} - o_{pj}) o_{pj}(1 - o_{pj}) \quad (18)$$

and error for a hidden unit, U_i, is

$$\delta_{pj} = o_{pj}(1 - o_{pj})\sum_k \delta_{pk} w_{kj} \tag{19}$$

The learning rate determines the changes in weights, which is proportional to $\partial E_p / \partial w$. The larger this rate, the larger the changes in weights. However, for rapid learning, a high learning rate may cause instability in the learning process. The momentum term is included in the generalised delta rule in order to increase the learning rate without causing any instability during the process. A reduction in the learning rate will take network to converge in longer time period.

Another factor that effects the convergence is the momentum term in the weight adaptation process. Momentum determines the effect of past weight changes on the current weight changes process. This process can be represented as

$$\Delta w_{ji}(n+1) = \eta(\delta_{pj} o_{pj}) + \alpha \Delta w_{ji}(n) \tag{20}$$

where;

n = presentation number

η = learning rate

α = momentum

The learning rate and momentum ranges between 0 and 1.

A backpropagation algorithm, with a sigmoid activation function, functions as follows (based on Lippmann, 1987 and Rumelhart *et al.*, 1992):

Start

Initialise the weights

Step 1 Present inputs and desired outputs.

$$X = (x_{i1}, x_{i2}, \dots\dots\dots, x_{in})$$

$$D = (d_{i1}, d_{i2}, \dots\dots\dots, d_{in})$$

Step 2 Calculate actual output.

$$o_{pj} = \frac{1}{1 + e^{-(\sum_i w_{ji} o_{pi} + \theta_j)}}$$

$$Y = (y_{i1}, y_{i2}, \dots\dots\dots, y_{in})$$

Step 3 Calculate error term for output nodes.

$$\delta_{pj} = f_j^{'}(net_{pj})(t_{pj} - o_{pj})$$

Step 4 Go back to hidden nodes and calculate error values.

$$\delta_{pj} = f_j^{'}(net_{pj}) \sum_k \delta_{pk} w_{jk}$$

Step 5 Adjust weights

$$\Delta w_{ji}(n+1) = \eta(\delta_{pj} o_{pj}) + \alpha \Delta w_{ji}(n)$$

Step 6 Until a desired level of error is achieved, go to Step 2. Once global minimum is achieved, show weights in the network.

3.2.3.7 Learning Issues and General Characteristics of Neural Networks

Unfortunately there is no rule of thumb that determines, for example the optimum learning algorithm or the activation function to be applied to a specific problem. It largely depends on experience, judgement and trial and error. Often it is not guaranteed that a network with given algorithm and training set will produce the desired output. As put forward by Wasserman (1993), there are many questions in regard to the learning process and they range from the applied training algorithm to the training sets. For example, the selected training algorithm may not be able to achieve the optimum weights for the desired mapping or ensure the network will train to the best set of values. Also there is the possibility that, with the gradient descent technique, the training converges into a local minimum (Feldman and Kingdon, 1995).

Even though the splitting of training data as training and validation helps in terms of when to stop the training, the process is still judgmental. Also with the application of gradient descent procedure there is always the danger that the training converges into a local minima prematurely (Feldman and Kingdon, 1995). One of the other issues from a practical perspective is the training time required to achieve such convergence. The training process can consume a lot of computer time on each training exercise. Different methods have been suggested to speed up the training time such as calculation of second derivative of the error term with respect to a weight in relation with the first derivative.

However, as Feldman and Kingdon (1995) pointed out more complex weight update rules must be applied to avoid cases of negative values in using the second derivative.

One other method mentioned by Feldman and Kingdon (1995) is the dynamic learning rate. This type of learning adjusts the learning rate (the value of η) in relation to each weight dynamically. Therefore the step size associated with each weight change is adjusted during the training.

However it is argued in the literature that some of the special characteristics of ANNs offset these negative issues. These are;

1. *Generalisation:* Generalisation is the ability of the network to solve new cases without being retrained. Once the training is complete the essential features of the problem are modelled and can be applied to a new data. The measure of generalisation is the difference between the two error rates of training data, e_1, and a new data, e_1, from the same distribution $| e_1 - e_2|$ (Wasserman, 1993).
2. *Non-linear Modelling*: The purpose of the training is to model the given problem, which in most the cases is non-linear and complex. This feature is important in this study and might provide advantage to ANNs in their modelling power of takeover predictions as some of the variables considered show non-linear characteristics. Morck *et al.*, (1988), for example, found that the relationship between the shareholdings of executive directors and shareholders' wealth maximisation is not linear. Their result can be an evidence of management entrenchment. Therefore it is possible that at high

levels of managerial ownership of companies, managers act to enhance their own interest rather than their wider shareholders. Stulz (1988) reported that the relationship between managerial shareholding and market value of the firm is non-linear. He identified such a critical value that if management shareholding is below this value then market value of the company increases otherwise it decreases.

3. *Tolerance to Noise:* Neural networks are tolerant to the noise in the data due to the properties of the hidden layers and the nodes in these layers, which distribute and store the information among the different neurons in the network.
4. *Tolerance to Missing Data:* In the event of missing data, the network completes the task by using partially available data. Missing data does not impede the network from retrieving the processed information from other neurons. For example, logistic regression does not produce an outcome variable if there is a missing value in one or any of the cases in the data while ANNs produces an outcome by using the partially available information.
5. *Parallelism and Connectionism*: The parallel connection in neurons allows the network to solve a problem simultaneously.
6. *The Performance Degradation:* Since the information is stored through the system rather than one part of the network, the performance of a network degrades slowly when damage or noise occurs.
7. *Adaptability:* As the new information is fed into the network, there is no need to retrain the network. The network adapts itself to the new environment.

3.3 Application of ANNs in Business, Finance and Accounting

Most of ANNs applications in finance and accounting are directed to empirical testing of the technique and benchmarking it against the more traditional modelling techniques of multivariate statistics. Feldman and Kingdon (1995) discuss some of ANNs applications in finance and highlight the increasing number of research papers published within the last five years.

Tam and Kiang (1992) applied ANNs to bank failure prediction in comparison with LDA, LR, *k* nearest neighbour and decision trees. The dataset used in the study comprise of 59 failed and 59 non-failed Texas banks. They have selected nineteen financial ratios and constructed two network structures. One with a two layer network (no hidden units) and the other one with three hidden layers with ten nodes. The results show that ANNs (three layer network) with ten nodes outperformed in its prediction all of the other classification techniques and the two layer network with no hidden nodes one year prior to bankruptcy. On the other hand, LDA has given better prediction results two years prior to bankruptcy. Overall their results suggest that ANNs offer an alternative/additional technique that can be utilised in monitoring banks in terms of their financial healthiness. However they also emphasise problems of deploying such a technique in terms computational efficiency (training time), the black box properties, as not being able to extract the relative importance of each input, and the network structure as there is no method to structure a network for a given classification task.

Yoon *et al.*, (1993) applied ANNs to prediction of stock price performance of 151 companies from 1989 and compared the results with DA. They have used four financial ratios as inputs and the companies are separated into two groups' being those whose stock price performed well or poorly depending on market value or total return. They further divided the dataset into estimation (76 firms) and hold-out samples (75 firms). They have employed back-propagation algorithm with sigmoid activation function. Different hidden units (0 to 9) with different hidden layers (2 to 4) have been tested. For a DA, the quadratic discriminant function was computed. The results show that the performance of ANN model increased with the increased hidden layers. While a two layer network with zero hidden unit successfully classified 65% of the hold-out sample, four layered network classified with 76% success. However the difference between the three and four layer network on the hold-out sample is the same even though the four layer network had a better mapping of the training sample. In comparison to DA, however, all of the constructed ANNs performed better both on the training and hold-out samples. DA only classified 63% of hold-out sample successfully. This is two percentage points lower than the two layer network. So their overall conclusion was that the multilayer ANNs has outperformed the quadratic DA model in forecasting stock price performance.

Curram and Mingers (1994) compared ANNs with LDA and decision tree induction rule. The comparison is performed on seven datasets from different fields. Four of these datasets were real and the remaining three were created artificially. Obviously the network structures were different for each of these seven datasets. The results of these seven datasets suggest that ANNs provided better classification results on the datasets where non-linearity existed. On the other hand, LDA performed better on the datasets that were linearly separable. One of the peculiar outcomes of this study is that they have found that the real datasets seem to be linearly separable compared with the artificial datasets. They also highlight the problems associated with ANNs such as training time and forming the network structures.

Luther (1996) applied ANNs to predict whether a firm that filed for bankruptcy under Chapter 11 will emerge reorganised through the bankruptcy process. It is stated in the study that those firms, which reorganise by the Chapter 11 process successfully, produce investment returns up to 50%. The data set is comprised of 73 reorganised and 31 liquidated firms. Luther also employed LR as a comparative modelling tool to ANNs. Thirteen financial ratios, one year prior to the bankruptcy filing, are used as inputs. Three layer, one hidden layer, neural net work model has been applied. The number of hidden units in the hidden layer was changed between 2 to 10 and performance of these different units was also compared. It is reported that 8 nodes in the hidden layer performed better than the other nodes. The training of ANNs is performed using Genetic Algorithm. The reported prediction accuracy of ANNs is higher than LR in all the different cut-off points. At 0.5 cut-off point, for example, the overall error rate for ANNs on the hold-out sample was 37% compared to 48% with LR. It is also shown that ANNs are less sensitive to cut-off points. The variation of error rate between different cut-off points with ANNs was only 12.5% (35% to 47.5%) point while it increases to 25% (38.5% to 63.5%) with LR. As a result Luther reported that as ANNs produced better prediction rates they could be used in bankruptcy prediction.

Miranda and Burgess (1997) applied ANNs to forecast market volatility changes implied in the transaction prices of the Ibex35 index options, which contains the 35 most liquid stocks that trade in the Spanish Stock Exchange. They have also applied classical statistical modelling, Linear Regression (LR) and Moving Average model

(MA), in order to compare the results. In the training of ANNs, they employed cross-validation selection procedures where the training data was split into training and validation sets. When the results of different models were comparedon the basis of root mean squared error and percentage of correct change predicted (correlation), ANNs and regression models performed similarly and compared with MA in terms of root mean squared error. However in predicting the correct change ANNs outperformed the regression model. Moreover they have carried further testing for better performance comparison between the models. The followed methodology is called comparison by forecast encompassing. This means that if one model is to be preferred to anotherone, the model one should explain what the other one could not explain. In order to establish which model encompasses the other one, first forecast error from model one is regressed on the forecast from model two and vice versa. If the resulting regression coefficient is not significant in model one but with model two then model one encompasses model two. This test simply estimates if model one's forecast can explain some of the forecast error of model two. On this test, it is reported that ANNs encompasseslinear regression and linear regression encompasses MA. ANNs is the only model whose forecast is not encompassed by any other models' forecast. The overall results in the study suggest that ANNs' forecast generally outperforms the forecasts of linear regression and moving average modelling.

Steiner and Wittkemper (1997) applied ANNs to portfolio optimisation. The model is based on the coherent market hypothesis (CMH) as a non-linear dynamic model of the capital markets. The nonlinear dynamic capital model, CMH, relaxes some of the assumptions of capital asset pricing model or arbitrage pricing theory such as rational investors and normal distribution of stock returns. Daily returns of highly liquid 31 shares that traded on the Frankfurt stock exchange are used to construct a simple market portfolio. The time period ranges from September 1991 to April 1994. First of all probabilistic ANNs are used to highlight the underlying stock return distribution from the previous two months stock returns. Then by using this information the CMH parameters are estimated. Based on the estimated parameters and the actual returns, a second general regression ANNs model is used to forecast the future returns of each stock relative to the market portfolio returns and each stock is ranked according to its performance. Two groups of stocks are established as high and low performance stocks. This process is repeated every day to update portfolios. The results show that a constructed portfolio outperforms the market portfolio by approximately 60%. Also, it is reported that the highest ranked 10 stocks show, on average, positive excess return. The overall results of the study conclude that as nonlinear modelling technique ANNs can be applied to excess stock return forecasting as well as being used as a tool to extract the underlying distributional properties of stock returns.

Serrano-Cinca (1997) applied ANNs (MLP) in corporate failures in 66 Spanish companies and compared the results of MLP with LDA and LR. First, he tried to observe the functioning of ANNs and its similarities with LDA. When a single layer perceptron with a linear transfer function in the output layer was applied to the whole sample the classification outcome was very similar to the LDA. Eight misclassifications with LDA compared to seven with ANN. In a similar manner a single layer perceptron with a logistic transfer function obtained similar results with LR. In this case four misclassifications compared to five with LR. Therefore he concluded that LDA and LR could be interpreted as a particular case of a single layer perceptron even though the three models are using different methods of estimating the parameters. In the second testing of these techniques Serrano-Cinca applied jackknife technique to LDA and ANNs. LDA produced nine misclassifications with 86.36% accuracy compared with four misclassifications in ANNs. In this case the network had four neurons in the hidden layer with a hyperbolic tangent as the transfer function and one neuron in the output layer with a linear transfer function.

The only application of ANNs to takeover predictions, so far, has been conducted by Sen *et al.*, (1995) on listed US firms in New York Stock Exchange. The study aims to compare the predictive and explanatory abilities of ANNs and logistic regression in corporate takeovers. The data collected over the period from 1980 to 1985 and 1984 data used to estimate the parameters and train the neural network. 1985 data on the other hand is used as a hold-out sample to test the predictive abilities of the models. Equal numbers of target and non-target companies (39 target and 39 non-target) are used in the estimation sample. The selected twelve company financial information is similar to the variables used by Palepu (1986).

Among the twelve variables used in the study only three of them (average excess return, average growth, and market to book ratio) were significant in the prediction of merger targets. They have also used the approach that is applied by Palepu (1986) to determine the probability cut-off point and the cut-off point of 0.6 is chosen as the intersect point between the probability distribution curve of targets and non-targets.

The neural network model with two hidden layers, four nodes in the first layer and two nodes in the second, and a hyperbolic tangent activation function was applied. Even though this network correctly classified 34 out of 39 targets and 37 out of 39 non-targets in the estimation sample, its prediction success on the hold-out sample was not so successful. Neural network predicted 28 out of 78 (36%) targets successfully compared with 23 (29%) logistic regression. On the non-target sub-sample however logistic regression had a better prediction rate (68%) than neural networks (63%). Overall neural networks correctly classified 61.2% compared with 66.4% of logistic regression. Moreover in order to show that the network can map the estimation sample with only one hidden layer, they trained a network with one hidden layer, 23 nodes in the hidden layer, to maximum learning. By doing this, the model is forced to over-fit to the training data. This network correctly classified 97% of the data but performed poorly (50%) on the hold-out sample.

Their conclusions suggest that ANNs performed similarly to logistic regression in the prediction of takeover targets even though none of the modelling techniques produced such results that could be used in the prediction of takeovers so that abnormal excess return can be obtained through their application. On the other hand as is shown in the study, neural networks provided better mathematical fit to the data.

3.4 Discussions and Summary

As can be seen from the brief and overall explanation of the techniques and the empirical application of them in business, finance and accounting, there are some similarities between LR and neural networks. Neural networks share a number of characteristics with traditional techniques. They include hierarchical structures, clustering, pattern association, and learning or training (Schalkoff, 1992). A two layered network with sigmoid activation function corresponds to LR. In a general sense, maximum likelihood estimation, where the unknown parameters are calculated by maximising the probability of obtaining the observed set of data through an iterative process, is very similar to the workings of neural computing, especially, when a sigmoid threshold function is used as the activation function. On the other hand, as can be seen from the characteristics of neural networks, there are also differences between them. For example, the missing data may result in LR not producing any score for the observation while the neural networks does. As reported by Sen *et al.*, (1995) for example, with noisy data neural networks model the data better than logistic regression.

In a more general comparison, the artificial neural network paradigms provide tools that are equivalent to those of classical classification and pattern recognition techniques. However as a distribution free method, ANNs are more robust to the cross sectional data and produce better solutions while multivariate statistical techniques, such as LDA, are subject to specific forms of distribution and assumptions. As the previous research suggests (McLeay, 1986; Ezzamel *et al.*, 1987; Barnes, 1987; So, 1987; Watson, 1990; Ezzamel and Mar-Molinero, 1990; Sudarsanam and Taffler, 1995), if the distribution of financial ratios deviates from normal distribution, the reliability of techniques that infer through this assumption will be questionable. In addition to normality, another critical assumption is that the group dispersion (variance-covariance) is equal across groups. Eisenbeis (1977) suggests that both assumptions are probably violated in most samples used in empirical research.

On one hand it is suggested or recommended that application of any statistical techniques such as MDA should be accompanied by trimming or transformation of variables until their distributions provide a reasonable approximation to normal distribution. Sudarsanam and Taffler (1985), in their study of industrial classification, found out that none of the 18 variables (financial ratios) were normally distributed hence they transformed these variables so that they approximate normality. Furthermore in their empirical analysis of proportionality condition of financial ratio components, Sudarsanam and Taffler (1995) found that most of the accounting ratios show loglinear relationship in their components and seem to suggest that most of these ratios may be lognormaly distributed. However, when they also examined whether this established relationship was constant over time, they found out that functional relationship of ratio component change over time resulting same type of transformation of same variables at different time points invalid. Their results on time variability of financial ratios support the findings of Mar-Molinero and Ezzamel (1989).

On the other hand though it is suggested by McLeay (1986) that rather than "*drawing inferences from trimmed means, or from re-expressed data, it would seem more straightforward to leave the data unadjusted and use a better-fitting model*"(p.209). Therefore it is only appropriate that in analysis and application of financial ratios one should try to construct models with statistical techniques, such as Logit or Probit analysis, that do not depend so much on normality assumption of accounting ratios. Also as mentioned above, one of the claimed advantages of ANNs is that they are free of multivariate normality. Consequently it is expected that ANNs will produce more robust results in finance and accounting applications. This expectation has been deducted as a result of the propositions and results of those studies that were summarised above.

As a result of above discussion that DA will not be applied in this study. It is quite obvious from the assumptions of DA that the technique is not suitable to the type of data that is intended to be utilised in this study. It is possible that DA could be applied to the financial ratios with application of variable transformation, but it would be highly inappropriate to apply it to the non-financial variables and eventually the combined variables as financial and non-financial. Binary variables of non-financial characteristics of companies would violate the normality assumption of DA such that it would bias the estimated error rates and lead to an incomparable classification outcome.

Chapter 4 – Financial Ratio Modelling and Results

4.1 Introduction

In this chapter ANNs and LR will be applied to takeover prediction by using publicly available information (financial ratios) of UK manufacturing companies. The aim is to assess, firstly whether it is possible to predict takeovers by employing these two techniques and secondly to see if ANN will provide a better prediction compared with LR modelling.

Ratios based on historical accounting information are often used for evaluating the financial condition and performance of firms and used as input in statistical models. Usage of financial ratios as performance measurement of companies comes from a very simple fact that companies operating in the general market or within specific industries are of different size. Ratios, therefore, are used to control for the effect of firm size as well as for industry wide factors.

Whether in corporate takeover predictions or profiling the characteristics of firms subject to takeovers, several studies (Simkowitz and Monroe, 1971; Stevens,1973; Singh, 1975; Belkaoui, 1978; Wansley and Lane, 1983; Dietrich and Sorensen, 1984; Rege, 1984; Walkling, 1985; Palepu, 1986; Barnes, 1990; Sen *et al,* 1995; Powell, 1997; and Barnes, 1998) have already used financial ratios to model takeover likelihood. The different ways to determine the likelihood for being takenover using past financial information is well documented.

Most of these studies have formulated successful prediction models for takeovers by using financial variables. The financial ratios that were employed by these studies as well the statistical methods used are briefly summarised below. The prediction results and successes of these studies will be reported in the summary section the end of the chapter.

Simkowitz and Monroe (1971) selected 24 ratios from seven different financial aspects of a firm and applied MDA. These are; growth, size, profitability, leverage, dividend policy, liquidity, and firm's stock. They have found out that seven of the 24 variables were significant. These reported significant ratios were; price-earning ratio, average dividend yield over three years, average growth rate of equity over three years, size, loss carry-over for tax, high volume of market activity, and average dividend yield to the equity over three years.

Stevens (1973) used 20 ratios under five categories that highlight the overall financial characteristics of firms. He selected these ratios from classes such as liquidity, profitability, leverage, activity, dividend policy, and price earnings. He reported that four ratios, EBIT / Sales (profitability of firms in relation to sales), Net Working Capital / Assets (liquidity measure), Sales / Assets, and Long-term liabilities / Assets (capital structure), found to contribute to the discriminatory ability of the models. He did not find that dividend pay out and price-earning ratio were significant. Some of his results contradict the conclusion of Simkowitz and Monroe (1971) for example that they found out that the price-earning ratio was a significant indicator.

Singh (1975) has constructed one of the earliest studies of takeover selection process in the UK. He used seventeen ratios from seven variable classes (Profitability, Change in Profitability, Growth, Liquidity, Gearing, Retention Ratio, and Size). He applied MDA and reported that Size, two-year average profitability, growth, two-year change in profitability, two-year average liquidity, two-year average gearing, and two-year average retention ratio were significant variables. The also reported that the model was able to classify 83% of the firms correctly.

Belkaoui (1978) employed 17 ratios under four different financial classes. He indicated that the ratios were selected on the basis of their popularity in the literature as well as their relevance to takeovers, and their distinction between liquid and non-liquid ratios. The classes are determined on the basis of their difference between liquid and non-liquid characteristics of companies. These were; Non-liquid Asset group (six ratios under this group), Liquid Asset to Total Asset group (four ratios), Liquid Asset to Current Debt group (three ratios), and Liquid Asset Turnover group (four ratios).

Dietrich and Sorensen (1984) on the other hand viewed takeovers (especially takeover decisions) as a capital acquisition decision and outlined that higher the value of expected net present value of cash flow from an acquisition the more attractive a possible takeover candidate become. In turn they concentrated on ratios (ten ratios) that directly or indirectly affect initial cash outlays and random future net cash flows. They have reported that Payout (dividends / earnings), Asset turnover (sales / total assets), Size (market value of the equity), and Trading volume were significant indicators of takeover likelihood.

Palepu (1986) applied a logistic regression technique to a group of ratios, which were derived from different mergers and acquisition hypotheses that were indicated as probable causes of takeovers in the literature. These hypotheses (Inefficient Management, Growth-Resource Imbalance, Industry Disturbance, Size, Market-to-Book, and Price-Earning) are further discussed in detail in the following section. He reported in one of the models that growth-resource dummy, growth, leverage, industry dummy, and size were determining factors in acquisition likelihood estimation.

Sen *et al.*, (1995) applied almost exactly the same hypotheses and ratios of Palepu (1986) to their takeover modelling of both logistic regression and artificial neural networks. They have reported that Average Excess Return, Average Sales Growth, and Market-to-Book values to be significant values. Compared to Palepu's (1986) study they did not find industry dummy to be significant.

Barnes (1990) selected nine basic financial ratios for his study. However he did not use these ratios directly in the modelling but averaged them with the relevant industry average (industry relative ratios). In the second stage he factor analysed them and found that five factors were explaining approximately 92% of the variance in the data. Hence five ratios that were most correlated with these factors used in the factor analysis. These ratios were; Quick Assets / Current Liabilities, Current Assets / Current Liabilities, Pre-tax Profit Margin, Net profit Margin, Return on Shareholders' Equity.

Powell (1997) modelled takeover likelihood of targets in the UK from the time period 1984 to 1991. He, like Sen *et al.*, (1995), constructed variables for acquisition characteristics of companies from similar hypothesis that were outlined by Palepu

(1986). He employed eight ratios (ROCE, Size, Tangible Fixed Assets / Total Assets, MTB, Operating Cash Flow / Total Assets, Average Change in Total Sales, Liquidity, and Leverage. However unlike above studies Powell (1995) did not apply the models to a hold-out sample to measure their classification/prediction abilities. He rather used them to see if there are differences between hostile and friendly takeovers as well as impact of change on acquisition characteristics of companies in time. He reported in the general model (consists of both contested and friendly takeovers) that Liquidity, measured as Cash and Marketable Securities / Total Assets, Size, measured as Log of Total Assets, and Market-to-Book ratio were significant variables. He further sub-sampled the data and divided into two separate time periods 1984–1987 and 1988-1991 to measure whether any significant changes could be observed in the characteristics of companies. He reported that the impact of firm characteristics on takeover likelihood changes over time. The model constructed with the data from period 1984-1987 produced only one significant variable (Size) while 1988-1991 model revealed three significant variables (Liquidity, Size and MTB).

Barnes (1998) also utilised similar hypotheses as outlined by Palepu (1986) in the construction of the variables. However rather than selecting a representative ratio for each hypothesis he formed a forty-two ratios and eliminated twenty-five of them after performing a correlation analysis. They were simply dropped from the analysis if they had a correlation coefficient greater than 0.65. This procedure reduced the list of ratios to seventeen and they were grouped under the hypothesis of Inefficient Management (seven ratios), Growth-Resource Imbalance (nine ratios), Size (one ratio), and Anticipatory Share Price Change (one ratio). He found that Profit Before Tax / Sales, Profit Before Tax / Shareholders' Equity, Total Remuneration / Sales, and Sales growth over last two years were significant in the Non-share price model.

Findings of these outlined studies above, such as Sen *et al.* (1995), Powell (1997), or Barnes (1998), are discussed in detail in the discussion and summary section. The purpose of discussing the results of these studies later in the discussion and summary section is to highlight as well as to measure the significance for the results of these studies alongside the results of this work.

Correspondingly, following the methodological principles of earlier studies outlined above, the aim of the first phase of this study is to identify and model those financial characteristics of two groups of companies as targets and non-targets by employing Logistic Regression (LR) and Artificial Neural Networks (ANNs). In the second part, these models will be applied on hold-out samples in order to find out whether it is possible to classify these companies into their correct groups. Furthermore the result of these models will be compared to determine if any particular modelling technique is better at classifying companies into correctly.

4.2 Methodological Issues

Zmijewski (1984) and Palepu (1986) have made critical examination of the methodology used by other studies in prediction studies and argued that there are three methodological problems that make those successful reports of previous studies (Simkowitz and Monroe, 1971; Stewens, 1973; Belkaoui, 1978; Dietrich and Sorensen, 1984) unreliable. These problems can be grouped under three headings;

i. The use of non-random, matched sampling in the model estimation eventually leads to a biased estimation of the model's parameters and through this acquisition likelihood;

ii. The use of non-random, matched sampling in the hold-out sample leads to a biased estimation of parameters affecting the models' predictive ability of the population;

iii. The use of arbitrary cut-off probabilities in prediction without specifying a decision context makes the prediction accuracy difficult to interpret.

It is argued that using matched sampling without appropriate modifications to the estimators will lead to estimation bias in the models' parameters. The estimation bias will result an overestimation of prediction results in the hold out sample. This bias occurs when the application of a choice-based sample is used in the estimation where the inclusion of targets is not random. The econometric justification of sucha selection is that the number of target firms (n_1) in the population (N) is so small that random selection would lead to a very small number of targets being included in the estimation sample. Such a sample would not convey the necessary information needed to form the model as it would be dominated by the non-targets (n_2) and lead to imprecise parameter estimates. Therefore, the information content of a sample would be greatly reduced. So, in order to improve the information content of the sample, equal target and non-target sampling is generally used in the derivation sample. It seems obvious that the main disadvantage of using such a choice-based sample in prediction modelling is the restriction on the model, when it is tested for other samples, that do not hold such an artificial composition (Bartley and Boardman, 1990).

However, Manski and Lerman (1977) and Manski and McFadden (1981) show that in a population such as takeovers or corporate failure, a choice-based sample provides more efficient estimates than a random sampling of the same size. Also, Cosslett (1981) and Maddala (1983) show that the bias introduced by a choice-based sampling for the Logit model is limited to the constant term only.

In a takeover prediction study such as this, the target information captured by the model that is constructed by random sampling would be weak, as there would be only few target companies in the estimation sample. As a result, a sample of one to one matching (choice-based sampling) is used to estimate model parameters in this study.

As can be seen in most of the previous studies, a one to one matched holdout sample is used to test the predictive ability of the models. Although this process does not affect the probabilities assigned to each individual firm, it nonetheless leads to misleading expected error rates. As one-to-one matching is a non-random sampling, the estimated error rates would not be generalisable to the population. In other words, the estimated error rates will not be equal to the expected prediction error rate in the population (Palepu, 1986).

It is also possible to utilise techniques such as bootstrapping or jackknifing rather than dividing the sample into estimation and hold-out groups. The jackknife technique, for example, requires that the estimation sample is formed with n-1 of the original cases and applied to the remaining one as a test case. Then the process is iterated until all of the cases are excluded from the estimation process one by one and used as a test case. Even though application of these methods enables all of the cases to be utilised in the estimation of the models, it is an extremely time consuming process (Serano-Cinca, 1997). In the case of this study for example it would take 300 iterations to complete the analysis for only one model. Furthermore these processes are not embedded in the software that is used this study. Therefore in the interest of time and practicality the sample is divided into two groups as estimation and hold-out. As presented later in the chapter, in this study the composition of the hold-out sample is formed to reflect its real composition as closely as possible.

In prediction studies the predictive classification of models of sample companies is based upon the cut-off probability. After the estimation of each firm's takeover probability, companies are classified into groups according to whether their individual acquisition probabilities are smaller or higher than the predefined cut-off value. In most of the prediction studies of takeovers and corporate failure alike researchers preferred to use an arbitrary cut-off probability of 0.50. As Palepu (1986) showed, a predefined cut-off probability of 0.50 is arbitrary and the appropriate cut-off probability should be determined by the decision context in which the model's predictions are to be used. In this study Palepu (1986) applied error-minimisation criteria to compute the optimum cut-off probability by using the estimated probabilities of companies in the estimation sample. However, this procedure assumes that the cost of classification errors of both groups (Type I and Type II) are the same (Hsieh, 1993). Although the error minimisation is the most popular measure that is used in the applied classification studies, certain types of misclassifications are more serious then others. As stated by Hand (1997), if the expected cost of misclassification can be quantified, an optimum decision surface can be selected to minimise this cost instead of error minimisation on the actual number. However as it will be explained later on that not only the overall predictive accuracy of the models will be presented but also a significance test (proportional chance criterion) will be performed. The aim of this test is to see whether the magnitude of these classifications is significant as well as to see that the results can not be attributable to simple chance or fluke.

4.3. Methodology, Data and Variables

The sample set for the UK companies is extracted from Acquisition Monthly's annual index for completed as well as failed UK public takeovers from January 1990 to December 1996. However, monthly issues of the journal had to be inspected in order to identify the bidding dates rather than the completion dates that have been supplied by the annual index reports. Financial data for the target companies is collected from the Financial Analysis Made Easy (FAME) and Datastream databases. The rationale behind the identification of bidding dates is obvious as the intention of a takeover becomes public, and share price data of targets are affected with this expected takeover. Also sometimes this pending process may take up quite some time as a result affecting the quality of data during the pending period.

It has also been decided to concentrate on manufacturing companies in order to exclude the specific industrial effects of financial services, oil and gas and utilities. By trying to select companies that are subject to more common environmental circumstances will enable us to form more stable models. Inclusion of financial sector characteristics, for example, might further delude this common environmental circumstances. This is especially important in the case of financial ratios due to the extreme variations between different industries (Mar-Molinero and Ezzamel, 1989; Sudarsanam and Taffler, 1985).

Moreover, in order to obtain the maximum information available, Public Limited Companies are targeted in the sample set. Therefore, the sample companies had to fulfil two criteria:

i. to be a fully quoted manufacturing firms and,

ii. to receive a bid during the time period outlined.

In the end, the list provided us with 103 target firms with available data in manufacturing from January 1990 to December 1996. At this point two types of sample division are followed. In the first instance 103 target companies are divided into two groups for each year, randomly to be included in the estimation and test samples. Each target company matched up with healthy companies from the same four-digit SIC code and one of these healthy companies is selected to be used in the estimation sample. The composition of this sampling is presented in Tables 4.2 and 4.3. Second, in order to test the predictive power of models in time, another estimation sample is formed with companies that were taken over from January 1990 to December 1994 and tested on the companies that were taken over during 1995 and 1996. The reason for establishing this second sample is the occurrence of population drift. Population distribution of 1990-94 might not be the same as of 1995 and 1996. This approach will measure the stability of the model in time since takeover characteristics of the market change in time. Powell (1997) tested whether takeover characteristics are consistent in time and reported that firm characteristics on takeover likelihood change over time. Some of the reasons why these characteristics change and models may not be stable over time are outlined by Barnes (1990) as: inflationary effects, technological reasons, changing accounting policies, and the change in the distributional cross-sectional parameters of financial ratios. The composition of this sampling is presented in Tables 4.2 and 4.3. In order to simplify these two approaches, the first one is called "Mixed Data" (MD) and the second one is "Time Data" (TD). Moreover, in order to see the industry relative performance of the firms, respective ratios of the companies are calculated as (Firm ratio/industry average) * 100 and called "Industry-Relative Data" (IRD). IRD is only computed for TD.

The reason for forming a data set as mixed data, where selected target firms from all years under consideration are included, is twofold. First, it was the preferred way of model estimation in some of the previous studies (Dietrich and Sorensen, 1984) and second, more importantly, it will be used as a benchmark for the TD model. Since the estimation sample as well as hold-out sample includes data from 1990 to 1996, it is expected that the prediction results of the MD model will be higher and a drop in the prediction power of TD model can be observed. It will not be attempted to construct IRD for MD because it is composition is already artificial and any significant result produced could not be taken as evidence that it is possible to predict takeovers. The purpose of having such a composition as MD is to see how much change, whether it is model significance or prediction result, will be observed compared with TD. A reality check is needed at this point. In actuality, it is only possible to make prediction for the coming event and not for what already has taken place.

The same approach was applied to the second year data to see the whether it is also possible to model takeover probabilities two years before the actual takeover bid. The term 'year' used in this context does not necessarily imply a calendar year but it actually represents accounting period. One year data means accounting information published just before the bid made public. Hence two year data means accounting data published one period prior to the one just before the bid made public. In order to differentiate these two data periods, acronym TD(-1) will be used for data one year prior to takeover bid and TD(-2) will be for the data two year prior.

As will be seen in the results, variables for the second year model were not significant (except LIQD) and the model itself is insignificant. This result indicated that, first of all, the variables contain more information about the companies' takeover likelihood one year prior to the takeover bid and posed a question of the possibility of using this information in the modelling process. Therefore, a different data set is formed by measuring the change in the variables, except Sales Growth (TUT) because it already measures the change in turnover, from one year to two years prior to the takeover bids and is called Variation Data (VD). VD, as in the case of IRD, is only calculated for TD.

It is also important to note here that the results of the IRD should be viewed with extreme caution as the industry averages have been calculated from two digit SIC codes rather than four digit as in the case of matching. The obvious reason for this is that these variables (except P/E and MBV, because they were to be extracted from Datastream) had to be calculated/extracted from the FAME and in some cases there is not a sufficient number of companies in some of the four digit SIC groups. The main problem with this type of averaging is that companies operate in such a wide array of other groups that using a two digit SIC simply can not be used as a benchmark against companies. Secondly, comparisons of company ratios against industry averages are problematic (McLeay, 1986; So, 1987). Sudarsanam and Taffler (1995) reported, for example, a huge variation in the functional relationship of ratio components between industries and recommended caution in the use of financial ratios in inter-industry comparisons. As mentioned by Mar-Molinero and Ezzamel (1989), and Lawrence (1982) that one of the main limitations of using industry averages is that product diversification hinders precise industry classifications and dilutes the averages. Also industry averages can be heavily influenced by extreme variations in the data that can not be observed especially if there is no remedial action, such as trimming, performed on individual company information that forms the industry data and average (Mar-Molinero and Ezzamel, 1989).

The reason for choosing this time period is to identify the characteristics of target companies before the expected stronger European integration that will be completed by the so-called monetary union. Also, selecting a seven year time period prohibits a longer time period for the financial data and some other effects such as changes in the market for mergers and takeovers (Powell, 1997 and, Jarrell and Bradley, 1980). As already mentioned above Powell (1997) examined whether takeover characteristics are consistent in time and found out that characteristics of companies on takeover likelihood change over time. His results suggest that in order to form stable models the estimation period should concentrate on a relatively short period of time where effects of changing economic environment do not alter the causes of takeovers. Mitchell and Mulherin's (1996) study confirms that takeovers do not occur evenly over time. Their result suggest that economic shocks at different times induce takeovers and restructuring in specific industries leading to increased acquisition activity. As a result it is important that the sample data should concentrate on a reasonable short period time so the effects of these shocks and their impact on industries can be captured. Rising stock prices along with the economic recovery started at the beginning of 1990s has resulted an increase in the takeover activity in the UK and the result can be observed in Figure 1 in chapter one.

The target sample set of 103 is comparable to the sample set of Gammie and Gammie (1996) who examined almost the same time period, 1990-93, and came up with 102 target firms (23 hostile an 79 non-hostile) and Weir (1997) who also used 1990-93 period with 94 target companies. Of these 94, 71 were non-hostile while 23 were hostile bids. However their data sets covered all sectors of the economy, not only manufacturing. The comparison of the sample with some other similar studies can also be made with Table-4.1 provided below.

Table 4.1 Sample and Technique Characteristics of the Some Major Studies

Study	**Sample Size Acquired/Non-acq.**	**Matching Criteria**	**Technique Applied**
Simkowitz and Monroe (1971)	23 to 25	Random	MDA
Stewans (1973)	40 to 40	Asset size	MDA
Belkaoui (1978)	25 to 25	Industry & Asset size	Dichotomous/DA
Rege (1984)	44 to 44	Industry	MDA
Dietrich and Sorensen (1984)	30 to 30	Industry	LR
Palepu (1986)	163 to 256	Random	LR
Bartley and Boardman (1990)	41 to 153	Industry	MDA
Barnes (1990)	92 to 92	Industry	MDA
Sen *et al.* (1995)	39 to 39	Random	LR/ANNs
Powell (1997)	411 to 532	Random	Multivariate Logit
Barnes (1998)	82 to 82	Industry & Size	Logit

Table 4.2. Composition of the estimation/derivation sample for the Mixed (MD) and Time (TD) data.

Year Acquired	MD	TD
Targets		
1990	9	16
1991	9	18
1992	7	15
1993	6	13
1994	8	16
1995	8	-
1996	5	-
Total Targets	52	78
Matched Firms	52	78
Total Sample	104	156

Table 4.3. Composition of the test/hold-out sample for the Mixed (MD) and Time (TD) data.

Year Acquired	MD	TD
Targets		
1990	7	-
1991	9	-
1992	8	-
1993	7	-
1994	8	-
1995	7	15
1996	5	10
Total Targets	51	25
Matched Firms	145	119
Total Sample	196	144

As can be seen in the previous tables, estimation samples are formed with one to one matching rather than random sampling. The reasons for this type of sampling method have already been explained above (Manski and Lerman, 1977; Manski and McFadden, 1981; Cosslett, 1981; Maddala, 1983). On the other hand, in the hold-out sample, this type of one to one matching is avoided. In the MD, target firms make up 26% (51/196) of the total hold-out sample and in the TD, they constitute 17.3% (25/144) of the total hold-out sample. Although these percentages do not exactly match population percentages, they are far from the 50-50 sampling composition. The FAME database on average consists of 500 manufacturing companies and if the 103 target companies are taken in total during the sampling period then 17 companies on average are taken over yearly which constitute 3.4% of the total. Also, it should be noted that 300 out of 500 manufacturing firms are included in the following analysis and this sample constitutes 60% of the total manufacturing population in the UK.

This matching is done on two bases. First, SIC; 4 digit SICs are used for matching in the estimation sample, and, second, financial year ends, in order to make sure that time differences are eliminated and to identify the main differences between the companies that are affected by the same external factors.

An analysis of financial ratios in decision models produces certain problems such as normality, missing data and outliers. Even though the techniques that are used in this study are free of the multivariate normality assumption for the independent variables and equal variance within each group, the issue of missing data and outliers and the best method for dealing with them remains. Missing observations can be problematic in analysis, although ANN is robust for missing values LR does not compute the cases where a missing value exists. SPSS provides several estimation techniques for replacing missing values. These are:

(1) Series mean. Replaces missing values with the mean for the entie series. (2) *Mean of nearby points*. Replaces missing values with the mean of valid surrounding values. The span of nearby points is the number of valid values above and below the missing value used to compute the mean. (3) *Median of nearby points.* Replaces missing values with the median of valid surrounding values. The span of nearby points is the number of valid values above and below the missing value used to compute the median. (4)*Linear interpolation.* Replaces missing values using a linear interpolation. The last valid value before the missing value and the first valid value after the missing value are used for the interpolation. If the first or last case in the series has a missing value, the missing value is not replaced. (5) *Linear trend at point.* Replaces missing values with the linear trend for that point. The existing series is regressed on an index variable scaled 1 to n. Missing values are replaced with their predicted values.

There is no specific guideline on this point and in order not to lose any cases in the data set, the missing values are replaced by the series mean.

Outliers, on the other hand, can be quite problematic in cross sectional analysis of companies from different industries. The presence of a few extreme values can have an important effect on summary statistics for example. There are different remedies for the treatment of outliers. These are listed by Kennedy *et al.*, (1992):

i no-adjustment,

ii trimming where a set percentage of observations are deleted,

iii winsorizing where the range of a variable is limited,

iv replacing the value of a variable with its relative rank,

v imposing a linear relationship within ranges (piecewise linear regression), and

(1) Series mean. Replaces missing values with the mean for the entire series. (2) *Mean of nearby points.* Replaces missing values with the mean of valid surrounding values. The span of nearby points is the number of valid values above and below the missing value used to compute the mean. (3) *Median of nearby points.* Replaces missing values with the median of valid surrounding values. The span of nearby points is the number of valid values above and below the missing value used to compute the median. (4) *Linear interpolation.* Replaces missing values using a linear interpolation. The last valid value before the missing value and the first valid value after the missing value are used for the interpolation. If the first or last case in the series has a missing value, the missing value is not replaced. (5) *Linear trend at point.* Replaces missing values with the linear trend for that point. The existing series is regressed on an index variable scaled 1 to n. Missing values are replaced with their predicted values.

There is no specific guideline on this point and in order not to lose any cases in the data set, the missing values are replaced by the series mean.

Outliers, on the other hand, can be quite problematic in cross sectional analysis of companies from different industries. The presence of a few extreme values can have an important effect on summary statistics for example. There are different remedies for the treatment of outliers. These are listed by Kennedy *et al.*, (1992):

i no-adjustment,

ii trimming where a set percentage of observations are deleted,

iii winsorizing where the range of a variable is limited,

iv replacing the value of a variable with its relative rank,

v imposing a linear relationship within ranges (piecewise linear regression), and vi fitting a polynomial model (nonlinear regression).

Also as shown by Kennedy *et al.*, (1992) application of the winsorizing and ranking procedures has resulted in a regression model that fits the data well and has a low level of prediction error. Mar-Molinero and Ezzamel (1989) have also highlighted the outliers' problem and stated that, according to the Tchebyschev inequality, if all the observations that are three standard deviations away from the mean are removed from the data set, the maximum error value that treating non-outliers as outliers is 0.11.

Therefore in this study observations located three standard deviations from the mean are accepted as outliers and replaced by the nearest observation. In this way the extreme observations are brought into a more reasonable range by winsorizing.

4.3.1 Variables

As it is true for every empirical analysis, success in the explanatory ability of any model is highly dependent and determined by the selection of the variables in the model. Unlike in corporate failure predictions where the lack of theoretical support leads researchers, for example Beaver (1966), Altman (1968), Ohlson (1980) and Taffler (1982), to select financial ratios due to their popularity and predictive success in the literature, the mergers and acquisitions literature provides hypotheses on the types of companies that are likely to become targets. Therefore, the financial variables applied in the models of the study are selected from the related mergers and acquisitions literature that has empirically tested these variables as predictors of the corporate takeovers (Rege 1984; Dietrich and Sorensen 1984; Palepu 1986; Sen *et al.* 1995; Powell 1997; Barnes

1998). These hypotheses and the ratios, as summarised below, are mostly used by Palepu (1986) and other researchers, such as Sen *et al.* (1995), Powell (1997), and Barnes (1998) in takeover prediction studies. Using these ratios will also provide the means for comparison of this research's results with the earlier studies.

Even though a single ratio is selected for each hypothesis, it is possible to employ groups of ratios as representative of each hypothesis. Barnes (1998), for example, used 18 ratios under four hypotheses because he stated that changing patterns and motivations of the market for corporate control make it unclear which variables are related to the probability of acquisitions. However, this type of model forming would lead to serious multicollinearity in the model resulting in bias in the model estimators.

4.3.2 Hypotheses

As mentioned earlier unlike in corporate failure prediction literature where there is a lack of theoretical support for which variables to employ in the model construction, merger and acquisition literature provides some underlying fundamentals or motives of takeovers. These fundamental reasons are frequently suggested in the acquisition or popular financial literature and consistently applied in the empirical analysis (Palepu, 1986; Sen *et al,* 1995; Powell, 1997; and Barnes, 1998), are employed in the variable selection process in this study as indicators of acquisition likelihood and their representative ratios are explained below. Obviously the six hypotheses or theories that are selected do not cover all the theories of mergers and acquisitions but provide a comprehensive coverage of most of the important indicators.

Hypothesis 1: Inefficient management hypothesis.

Inefficient management is replaced by more efficient management by means of takeovers (Palepu 1986). Jensen and Ruback (1983) viewed the takeover market as where managers compete for corporate resources. Hence, firms with inefficient management teams are replaced by other management teams for better utilisation of corporate resources. Those firms with under-utilised resources are more likely to be acquired and their management teams will be replaced by more skilled teams. The market for corporate control was first put forward by Manne (1965), providing the theoretical foundations for most of the research in takeovers afterwards. According to Manne (1965), a fundamental premise underlying the market for corporate control is the existence of a strong relationship between corporate managerial efficiency and the market price of shares of that company. The stock market provides an objective evaluation of management performance through the value it places on a firm's equity. Therefore, inefficient management would be evaluated on the market by a low equity price that will eventually create an incentive for others to take control of the firm for better utilisation of the firm's assets. Obviously, a high share price of an equity not only deters other firms from bidding because of its cost, but is also a sign that the market is quite happy with the way assets are utilised by the incumbent management.

Agency Theory argues that the threat of a takeover provides a permanent mechanism to monitor management performance and discourages the pursuit of management interests at the expense of shareholders. The nexus of contracts is at the centre of agency theory and it focuses on the inherent conflict between the interests of shareholders and the interests of those who run and work for the firm (Gerald and Stout, 1992). Walkling and

Long (1984) provide some insights into this conflict by analysing cash tender offers and management resistance to a takeover. They concluded that management resistance to a takeover is significantly affected by potential wealth changes of officers and directors. Wealth change of directors and officers is defined as summation of offer-induced changes in share and option wealth.

In order to measure the management efficiency or inefficiency Return on Shareholders' Fund (RSHF) has been selected as an indicator. RSHF was also the preferred variable for this hypothesis in Palepu's (1986) study while Powell (1997) used the accounting rate of return (ROCE). Some of the previous studies, such as Barnes (1990), have used the profit figure, such as pre-tax profit margins, as a measure of a company's performance. As can be seen these ratios are to cover the firm's managerial performance.

Hypothesis 2: Growth-Resource Imbalance.

Firms with high growth and low resources or low growth and high resources are likely to be acquired (Cosh *et al.*, 1980; Levine and Aaronovitch, 1981; Palepu, 1986; Powell, 1997; Barnes, 1998). Growth-resource mismatch dummy (GRDUMMY) along with Sales Growth (TUT), Liquidity (LIQD), and Gearing (GEAR), are used to test the hypothesis. GRDUMMY is estimated from Liquidity, Gearing and Sales Growth. GRDUMMY is assigned a value of one for combinations of high growth-low liquidity-high gearing or low growth-high liquidity-low gearing. Otherwise it is assigned as zero. These variables have been utilised by other studies besides Palepu (1986) (Ambrose and Magginson, 1992; Powell, 1997; Barnes, 1998) and appears it to be an important variable. Sudarsanam *et al.*, (1996) estimated a different growth-resource imbalance variable that takes bidders' as well as targets' growth-resource characteristics into account. However in this study growth-resource imbalance is estimated in a similar fashion to Palepu (1986), Sen *et al,* (1995) and Powell (1997) in order to be in line with these earlier takeover likelihood estimation studies. The reason why growth-resource mismatch might be significant is that high growth firms with low resources might be acquired or targeted by companies with the opposite growth-resource imbalance or vice versa (Powell, 1997).

Hypothesis 3: Proportion of Fixed Assets.

The hypothesis states that the greater the percentage of tangible fixed assets in a firm's total asset structure, the higher the likelihood that the firm will become a target. Ambrose and Megginson (1992) tested the hypothesis and found it to be a significant indicator. The rationale behind this assumption is that fixed assets are easy to value compared with intangible assets and can be used for debt financing to finance the takeover (Stulz and Johnson, 1985). Second, asset rich firms in declining industries could attract attention as a method of restructuring the firm (Ambrose and Megginson, 1992). The proportion of fixed assets (PROPFIX) is estimated by Fixed Assets divided by Total Assets (FA plus CA).

Hypothesis 4: Size.

Smaller companies are more likely to be acquired, due to size related transaction costs (Dietrich and Sorensen, 1984; Conn, 1985). The hypothesis considers the cost of acquisition and integration of targets into the acquiring firm. These costs are likely to increase with the target size. Some of the previous studies (Levine and Aaronovitch, 1981; Palepu, 1986) found size to be a significant variable.

In order to measure this size effect Current Liabilities are deducted from Total Assets (FA + CA) and transformed logarithmically.

Hypothesis 5: Price-Earning Ratio (P/E).

The rationale behind the P/E hypothesis is that bidders with high P/E ratios acquire low P/E firms to acquire an immediate capital gain based on the belief that the stock market values the earnings of the combination at the higher P/E ratio of the acquirer. This calculation is based on the expectation that P/E ratio of the bidder would remain the same after the merger (Hutchinson, 1995). No special remedies are applied to the negative values of P/E other than winsorizing that was explained earlier in this chapter.

Hypothesis 6: Firm Undervaluation (Market to Book Ratio (MBV)).

Firms with low market values compared with their book values are likely to be targets, since a low market-to-book value ratio is an indicator of cheapness. Palepu (1986) views these variables with suspicion since the book value of a firm does not necessarily reflect the replacement value of its assets. The appropriate measure for this value should be Tobin's q. The lower the Tobin's q, the higher the probability of receiving a bid. Tobin's q is also an indicator of management inefficiency. However in this study MBV is used for this hypothesis because neither replacement costs nor current costs are consistently produced by UK firms during the period (Powell, 1997).

The above hypotheses and their respected variables as well as their expected signs are summarised below at Table 4.4. Furthermore the more detailed definitions and estimation of the variables can be found in Appendix F.

Table 4.4 Takeover Hypothesis, Variables and Expected Signs

Hypothesis	Variables	Expected sign
1.Inefficient Management	Return on Shareholders' Fund (RSHF)	-
2.Growth-Resources Imbalance	Sales Growth (TUT), Liquidity (LIQD), Gearing (GEAR), and High-Growth Resource-Poor or Low-Growth Resource-Rich (GRDUMMY)	n.a.
3.Fixed Asset	Proportion of Fixed Assets to Total Assets	+
4.LOG(Size)	Total Assets (FA+CA)-Current Liabilities	-
5.Price-Earnings (PE)	Price-Earnings Ratio (PE)	-
6.Market to Book	Market to Book Value	-

4.4. Univariate Data Analysis Results

Tables 4.5, 4.6, 4.7 and 4.8 provide the summary of descriptive statistics for the MD, TD(-1), TD(-2), and IRD. As can be seen along with mean and *t*-values, median and Wilcoxon non-parametric test for median differences are also provided. Even though no analysis has been done to test the proportionality condition of financial ratio components of the variables used in this study, Sudarsanam and Taffler (1995) reported that accounting ratios do not satisfy the proportionality condition and for many ratios lognormal distribution may provide a better fit to underlying distribution. Their conclusion means that median rather than mean might provide a better statistic for the central tendency.

Further to univariate analysis correlation matrixes are supplied in Appendix C.

Table 4.5. t-statistics: Mixed data (MD)

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
RSHF	2.1929	27.9581	11.555	20.780	-3.55***	-3.74***
TUT	1.3452	12.8958	0.820	9.060	-3.29***	-3.29***
GEAR	82.9484	52.2608	64.025	40.225	2.33**	2.00**
LIQD	0.9640	1.0280	0.938	0.993	-1.47	-2.17**
SIZE	72029	102161	28770	24701	-1.25	0.003
PROPFIX	0.3913	0.3706	0.363	0.351	0.56	0.35
P/E	7.7786	17.3355	9.760	10.368	-1.26	-1.01
MBV	9.8327	14.2135	6.200	11.496	-2.19**	-2.49**
GRDUMMY	0.3076	0.3461	0.000	0.000	-0.40	-0.41

Notes:
1***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
2 Observations (52 Targets and 52 Non-targets)

As can be seen in Table 4.5, the mean and median differences in RSHF and TUT are statistically significant at the 1% level. Also GEAR and MBV are significant at the 5%

level. These indicate that non-target companies have higher sales growth and return on their shareholders' fund than the target firms. These firms, as a result, increased their MBV and their share price. Hence non-target firms have higher P/E and MBV. Results are consistent with the hypothesis outlined above that target companies are under-performers compared with matched ones. On the other hand, none of the other variables are statistically significant. Nevertheless, it is obvious that target firms have higher gearing, are less liquid and are smaller. In terms of fixed assets, however, there is no real difference. It is interesting to note that, even though the mean or median difference is not statistically significant, GRDUMMY for non-target companies is higher. This result is not expected as it implies that there are more growth-resource mismatch companies in the non-target group than target sample. The median test also affirms the *t*-statistic results except in LIQD that becomes significant at the 10% level. In addition median statistic shows that contrary to mean, non-target companies are actually smaller than target ones. However this is not statistically significant.

Table 4.6. t-statistics: Time data (TD -1)

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
RSHF	5.4021	23.1870	10.330	19.400	-3.22***	-3.29***
TUT	0.1400	11.0300	-1.280	7.020	-3.35***	-3.32***
GEAR	78.4541	57.4397	64.025	54.155	2.61**	1.97**
LIQD	0.9571	1.0530	0.968	1.005	-2.54**	-2.15**
SIZE	50322	110362	25527	22242	-2.21**	-0.14
PROPFIX	0.3819	0.3595	0.364	0.338	0.86	0.91
P/E	5.2311	13.9021	9.051	10.565	-1.67*	-1.21
MBV	9.8258	12.7729	7.500	8.108	-1.94*	-1.84*
GRDUMMY	0.3200	0.3070	0.000	0.000	0.17	0.16

Notes:
1***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
2 Observations (78 Targets and 78 Non-targets)

Table 4.6 shows that the differences in means are statistically significant in all of the variables except in the cases of PROFIX and GRDUMMY. RSHF and TUT are significant at the 1% level. GEAR, LIQD, and SIZE are significant at the 5% level and P/E and MBV are significant at the 10% level. Target firms make a significantly lower return on their shareholders' fund than the matched ones and their market values as a result are lower in comparison. The significant SIZE variable indicates that target companies are smaller than their counterparts in the matched sample. However the median statistic of size indicates otherwise. Median shows that, like in MD data sample, non-target companies are actually smaller in size even though the difference is not statistically significant. Also, targets show significantly lower liquidity and growth as well as higher gearing in line with the growth-resources imbalance hypothesis This can also be observed from GRDUMMY, contrary to MD data sample, GRDUMMY for target companies is higher. The overall result from these statistics show that target companies are under-achievers compared with those which had not received any takeover bid during the time period examined. These results are in line with the hypothesis outlined above.

The overall results for MD and TD suggest that the target companies are inefficiently run, had lower profitability, and as a result under-performed. This resulted in them either being taken over or receiving a bid so that their assets can be utilised better under a new management.

It is noteworthy to mention that some of the statistics have become significant with the time data set that includes more companies namely 156 compared with 104 in the mixed data set. This is encouraging due to the fact that these results support almost all of the hypotheses postulated above.

Table 4.7. t-statistics: Industry relative statistics for the time data (IRD)

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
RSHF	-0.3259	1.5646	0.647	1.454	-2.66***	-3.48***
TUT	-0.0013	0.3594	-0.075	0.362	-1.02	-2.55**
GEAR	1.0771	0.7550	0.874	0.525	2.80***	2.28**
LIQD	-5.9752	1.8980	3.424	10.056	-1.14	-1.17
SIZE	247	483	132	104	-2.18**	0.07
PROPFIX	0.8797	0.8265	0.848	0.824	1.02	0.78
P/E	0.3761	0.9443	0.510	0.775	-2.04**	-2.11**
MBV	0.6481	0.9389	0.515	0.803	-2.54**	-2.15**
GRDUMMY	0.2692	0.2820	0.000	0.000	-0.17	-0.17

Notes:
1***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
2 Observations (78 Targets and 78 Non-targets)

As mentioned earlier, above results of the industry relative data should be viewed with caution. Table 4.7 shows that RSHF and GEAR are significant at the 1% level. SIZE, P/E and MBV, on the other hand, are significant at the 5% level. The results are mostly in line with the TD data set, but at a lesser significant level. The only notable difference is that TUT does not seem to be significant compared to the other data sets with *t*-statistic but it becomes significant when median is considered. This means that when TUT is adjusted with the industry average the significant differences that were observed in the earlier data sets disappear. However in line with TD, target companies in relation to their industry are still significantly smaller with high gearing and provide lower return on their shareholders. A direct result of this can also be observed with their significantly lower P/E and MBV ratios. Like in TD (-1) data sample, SIZE becomes insignificant and non-targets companies are smaller in comparison to their counterparts in target sub-population.

Table 4.8. t-statistics: Time data (TD -2). Two years prior to takeover.

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
RSHF	17.5117	25.0624	17.230	20.235	-1.90*	-1.45
TUT	6.4680	10.0708	7.720	8.000	-0.96	-0.55
GEAR	66.3651	67.8638	53.455	54.515	-0.14	0.35
LIQD	0.9380	1.0029	0.952	0.975	-2.21**	-1.77*
SIZE	59693	55483	27056	20477	0.29	0.94
PROPFIX	0.3756	0.4773	0.359	0.347	-0.97	-0.09
P/E	11.3584	12.5761	10.249	9.157	-0.39	1.35
MBV	11.0238	13.0740	8.880	8.745	-1.14	-1.07
GRDUMMY	0.2940	0.3460	0.000	0.000	-0.66	-0.68

Notes:
1***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
2 Observations (78 Targets and 78 Non-targets)

As expected the quality of the information in the data deteriorated two years prior to the takeover bid compared with the one year prior data. Furthermore the signs of some of the variables did not come up as expected. GEAR, SIZE and PROFIX have opposite signs to the outlined hypothesis as well as MD and TD data sets. The results are quite interesting, as it shows a drastic change of target characteristics from two years prior to takeover to one year. The figures above indicate that target companies are actually bigger than non-target companies with less fixed assets in their books. Moreover they are less geared and there are more resource imbalance companies in the non-target group compared with the target sample. However their lower return on their shareholders' fund and lower P/E and MBV ratios suggest that they were already performing badly compared with their non-target counterparts even though the differences are not as significant as one year prior to takeover. LIQD seems to be the only variable that is significant with both parametric and non-parametric statistics indicating that target companies are less liquid compared to non-target companies.

Table 4.9. t-statistics: Variation Data (VD)

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
RSHF	-2.7365	0.1579	-0.179	-0.108	-1.78*	-1.55
TUT	0.1400	11.0300	-1.280	7.020	-3.35***	-3.32***
GEAR	1.0752	0.4706	0.120	-0.054	0.90	2.60***
LIQD	0.0256	0.1282	-0.011	0.010	-1.16	-0.96
SIZE	0.0082	0.8599	-0.001	0.077	-2.20**	-3.48***
PROPFIX	0.0222	-0.0188	0.001	-0.025	1.17	1.57
P/E	-0.3833	-0.4402	-0.144	0.005	0.09	-2.09**
MBV	0.0391	0.2972	-0.050	-0.008	-1.51	-1.02
GRDUMMY	0.0256	-0.0384	0.000	0.000	0.72	0.72

Notes:
1***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
2 Observations (78 Targets and 78 Non-targets)

Results in Table 4.9 show that the target companies experienced a significantly lower growth compared with non-target companies. TUT is significant at the 1% level at both in parametric and non-parametric statistics. SIZE is significant at the 5% level as measured by t-statistic and at the 1% by Wilcoxon Z. RSHF is significant at the 10% level. The negative return on shareholders' funds along with positive changes in TUT and SIZE suggest that even though the target companies increased their size and growth from two year to one year prior to takeover, management could not translate these changes into return for their shareholders. Even though these changes are significantly lower compared with non-target ones it still indicates a growth on the target companies. As a result the target companies had a negative change in their P/E ratios and this change is significant according to Wilcoxon Z statistic. Apart from those three significant variables, the P/E ratio and market value of the target company deteriorated in proportion to the non-targets. On the other hand, gearing increased comparatively more for the target companies. In the mean time PROFIX for target companies increased while non-target companies experienced a negative growth.

4.5. Multivariate Models and Results

The logistic regression estimation of the five models is presented below. Table 4.10 shows the estimates of one year prior to the takeover bid while Table 4.14 presents the estimates for the two years prior to takeover bid. Unfortunately, there is no way of determining the individual weights of each variable in ANN models to compare it with LR models, therefore each ANN model's classification tables and percentages for estimation samples are displayed under each table where the LR models are presented. Although these comparisons of classification tables do not show weights of variables, they provide a guideline of each ANN's mathematical fit to the estimation data. Also, it should be emphasised that this study is concerned more with the prediction powers of each model than the model's estimators.

All the computing was performed on standard Pentium processor. NEUframe Version 3 from Neural Computer Science for development of the standard back propagation ANNs. There are many other commercially available packages. However this particular software seems to be the most popular one both within academia and industry. Alici (1996) used this particular software successfully in his study on bankruptcy prediction applications. On the other hand, LR analysis was performed on SAS (v.8.2).

The neural network models for each data group were obtained after many trials of different nodes and hidden layers. It is important to perform multiple training runs in order to eliminate the possibility that the network might be caught at local minima. Although it has been shown (Cybenko, 1989 and Hornik *et al.*, 1989) that a single hidden layer network structure with sigmoid activation function can approximate any function with arbitrary accuracy, it was found in this study, after many trials of single hidden layer networks, that a two hidden layers' network provided a higher generalisation on the hold-out sample. This might be as a result of lesser hidden units and weight (nodes) being required in two layer network which sometimes improves generalisation. This of course further complicated the training process that is usually time consuming. Many combinations of hidden layers and nodes in these hidden layers were constructed before an optimum model was determined. The sigmoid activation function with learning rate of 0.1 and momentum value of 0.8 is used in this study. The learning and momentum rates define the speed and accuracy with which the backpropagation algorithm converges on the optimum solution through error minimization. Higher values for these rates induce speeder convergence but run the risk of overshooting the optimum solution. Taresenko (1998) indicates that a typical learning rate should be between 0.01 and 0.1 with a momentum value between 0.5 and 0.99.

Even though the training process/algorithm is explained in detail in the previous chapter, it is worthwhile to summarize it here for better understanding of its workings. First of all before the training the weights between neurons are randomised. During the training, input data (variables) are presented to the network one case at a time. The network then tries to predict whether it is a target or non-target and compares it with the correct answer. It then records the error, adjusts the weights between the neurons and repeats the process. Also the network is specified to stop at intervals to check its progress on the test sample. This iteration process goes on until further training fails to improve the network (Gately, 1996).

The summary of the training process followed in this study can be summarised as:

i. The data has been partitioned into training, validation and test sets.
ii. The network has been trained until the stopping criterion has been met.
iii. The optimum network has been selected and tested on the test set.

Since this research is primarily concerned with the predictive power of the models, the produced prediction results or results of the prediction tables of the hold-out sample should be measured accurately. As will be seen, the prediction results of LR and ANNs models are presented in a classification matrix form which is similar to SPSS's classification tables. It is a common practice to use the overall predictive accuracy of the models presented and accordingly this figure is used to measure the success rate of classifications. As pointed out by Hair *et al.*, (1995), it is essential that the magnitude of the overall predictive accuracy is greater than the maximum classification that could be attributable to change. Accurate prediction of group membership or predictive efficiency can be measured in several ways. Some of these measures are listed by Menard (1995) as phi, Goodman and Kruskal's gamma, kappa, the contingency coefficient, Pearson's r and the odds ratio. However in this study t-statistic will be used for its simplicity. In unequal group sizes, like the TD hold-out sample where there are 119 non-targets and 25 targets, the proportional chance criterion is 71.3% and computed as (Hair *et al.*, 1995):

$$c = p^2 + (1 - p)^2$$

where

c = proportional chance criteron

p = proportion of companies in group 1

1-p = proportion of companies in group 2.

On the other hand, the proportional chance criterion for MD is 61.5%. Therefore, a prediction accuracy of above 71.3% for TD and 61.5% for MD would be acceptable in order to accept the models' classification rates as successful.

The statistical test for the classification accuracy of the model when compared with a chance model is *t*- statistic and computed as (Barnes, 1998):

$$t = z - c / [\, z(1 - z)/\mathrm{n}]^{1/2}$$

Where

z = the proportion of companies correctly classified

n = total sample size.

The *t*-statistic will only be provided if and when the overall prediction rates of the models exceed the proportional chance criterion, because if the classification accuracy were not higher than the proportional chance criteria, the *t*-statistic provided would be misleading. As put by Hair *et al*., (1995):

"...... if the percentage of correct classifications is significantly larger than would be expected by chance, an attempt can be made to interpret the discriminant functions in the hope of developing group profiles. However, if the classification accuracy is no greater than can be expected by chance, whatever structural differences appear to exist merit little or no interpretation; that is, differences in group profiles would provide no meaningful information for identifying group membership."

In a decision context, two types of error occur: Type I and Type II. Type II classifies target companies as non-targets. On the other hand a Type I error classifies non-target companies as targets. The overall objective of prediction is to minimise both types of errors. The method of determining the error-minimisation, for example, depends on this assumption (Hsieh, 1993). However, in takeover predictions a Type I error is more costly than the Type II error. Although a Type II error means an opportunity lost for not investing in those prospective target companies for excess return, a Type I error means that investment will be tied up in companies which will not become a takeover target within a year. Altman and Eisenbeis (19978), and Tam and Kiang (1992), mentioned in bankruptcy prediction studies for example that priory probabilities of failure and the relative cost of Type I and Type II predictions are not equal.

If any of the prediction results are higher than what could be expected from a proportional chance criterion the results would become significant and merits further analysis on the Type I or Type II error. If they are not significant or higher than would be expected proportional chance criterion than the results would not merit any further interpretation (Hair *et al.,* 1995). Second, it is possible that by increasing the cut-off probability to higher percentages would certainly reduce misclassification of non-target companies in the event of modelling takeover likelihood. Barnes (1998), for instance, experimented with different cut-off probabilities in this way in order to classify more targets correctly with fewer non-targets classed as targets. Even though this is a viable option with LR, it is not possible to manipulate cut-off score of ANNs at least with the software that is utilised in this study. The Neural network software that is used in this study only produces two possible outputs as targets and non-targets, and it is not possible to obtain individual scores or probabilities as such so that cut-off point can be manipulated. Hence the black box argument of ANNs. As a result, in order to facilitate a comparison of the methods equal expected misclassification costs are implicitly being assumed in prediction rates. Therefore the error minimisation procedure will be applied to test the statistical significance of overall classification results in order to be able compare the outcome of these two techniques.

However as a supplemental measure percentages of target companies within the total predicted targets will also be displayed for comparison.

4.5.1 Estimated Models

Table 4.10. Logistic Regression Models for MD, TD and IRD. One year prior to takeover bid.

Variables	MD	TD	IRD
RSHF	-0.0459	-0.00968	-0.1512
	(5.474)**	(1.607)	(2.326)
TUT	-0.0518	-0.0254	0.00796
	(7.646)***	(5.705)**	(0.0087)
GEAR	0.00917	0.0026	0.4629
	(3.707)*	(0.5118)	(3.633)*
LIQD	-3.2318	-2.2717	-0.00382
	(5.919)**	(5.429)**	(0.744)
LOG(SIZE)	0.225	-0.0781	-0.2159
	(0.215)	(0.062)	(0.652)
PROPFIX	-0.2918	0.3577	0.605
	(0.040)	(0.095)	(1.314)
P/E	-0.0088	-0.00826	-0.1037
	(1.387)	(1.698)	(0.984)
MBV	-0.0467	-0.0188	-0.4313
	(3.309)*	(0.994)	(2.307)
GRDUMMY	0.1751	0.1549	-0.0626
	(0.0925)	(0.152)	(0.0269)
Constant	3.5630	2.8456	0.0832
	(2.051))	(3.011)*	(0.0148)
N	104	156	1564
Model chi-square	31.6463***	26.1999***	18.354**
Likelihood Ratio Index	0.2925	0.1364	0.1000
Correct classification %			
Target	69.23%	69.23%	58.97%
Non-Target	80.77%	57.69%	67.95%
Overall	75.00%	63.46%	63.46%

Notes:

1 ***Significant at the 1% level, 2 tail test

**Significant at the 5% level, 2 tail test

*Significant at the 10% level, 2 tail test

2 Figures in the parentheses are Wald statistics.

3 The likelihood ratio index is defined as (1-(log likelihood at convergence/log likelihood at zero)). It is similar to the R^2 statistic of multiple regression and gives an indication of the Logit model's explanatory power.

The Logistic regression results of for MD, TD and IRD are presented above in Table 4.10. A positive sign on a parameter/coefficient indicates that an increase in the associated variable increases the likelihood of takeover. On the other hand a negative sign decreases the likelihood.

MD and TD are significant at the 1% level while IRD is significant at the 5% level. These significance levels indicate that the above three models provide a statistically significant explanation of a firm's acquisition probability. Likelihood ratio indices are lower in TD and IRD compared with MD. The likelihood ratio index provides an indication of the overall explanatory power of the models and as explained above is similar to the R^2 statistic of multiple regression. So that at 0.2925, the MD provides a better explanation of a firm's acquisition probability. TD has a lower likelihood ratio index, 0.1364, that indicates a lower proportional reduction in error rate. The index is even lower with the IRD.

In the MD model, the variables RSHF, TUT, GEAR, LIQD and MBV are statistically significant. The coefficient signs of these variables indicate that companies with low turnover trends, low liquidity, low return on their shareholders' fund and low market to book ratio have a higher probability of becoming takeover targets. On the other hand target companies are geared significantly higher than non-target companies. However SIZE, PROFIX, P/E and GRDUMMY do not have statistically significant coefficients indicating that they do not contribute significantly to the acquisition likelihood. The results are in line with the univariate analysis and parameters' signs are as expected.

In the TD(-1) model on the other hand, the variables TUT and LIQD are statistically significant. Unlike in the MD model, RSHF, GEAR and MBV are not significant. The coefficient of these two variables are negative indicating that the companies, which have low liquidity and are slow in growth have higher probability of becoming targets during the examined period. Hence, the target companies in the estimation sample are characterised by lower growth (TUT) and liquidity (LIQD). Although the other variables are not statistically significant, their coefficients have expected signs. Other than GEAR, PROPFIX and GRDUMMY, all the other variables have negative coefficients as expected. The coefficients of the variables in the TD model indicate that target companies are less profitable, have a negative sign for RSHF, are highly geared, with a positive GEAR, and proportionally have more fixed assets in their asset structure than the non-target companies in the sample. Unlike Ambrose and Megginson's (1992) finding of significant proportional fixed assets, PROPFIX is not significant in the TD model, although the targets had higher fixed assets proportional to total assets. Also compared with Palepu's results (1986) GRDUMMY is not statistically significant in this sample.

Finally in IRD, GEAR is statistically significant at the 10% level while none of the other variables seem to be significant. This indicated that when the company characteristics are adjusted with their industry average ratios the differences between target and non-target companies disappear. Although the model is significant at the 5% level, its likelihood ratio index and its classification rate are smaller than for the TD and MD. As it is mentioned above the results of the IRD model should be viewed with caution in line with the arguments explained before.

Below tables present the classification rates of ANN models on the derivation/estimation samples of the above data samples. The classification rates provide mathematical fit of the network models on the estimation samples. They are useful in comparing to LR classification on the estimation samples.

Table 4.11. ANNs classification table for MD. 8-5 nodes in the hidden layers.

	Predicted		
Observed	Non-target	Target	%
Non-target	45	7	86.53%
Target	15	37	71.15%
		Overall	78.84%

In the case of MD, a two-hidden-layer network with 8 nodes in the first hidden layer and 5 nodes in the second hidden layer proved to be the best model which generalises well on the test sample. It is quite possible to train the network to reach better classification rates by training the network to an error limit of 0.05 that almost classifies almost all the cases correctly. Further training of this model, for example produced a classification of 49 non-targets and 47 targets successfully with an overall classification rate of 92.3%. But the model lost its generalisation ability and performed poorly on the hold-out sample classifying 32 targets and 87 non-targets correctly with a 60.71% overall classification rate. In comparison with the LR model for MD, the ANN model had a slightly higher classification rate overall and on both the target and the non-target companies. It successfully classified 45 of the 52 non-targets with an 86.53% classification rate compared to 37 of the 52 targets with a 71.15% classification rate. Overall classification rate is 78.8%.

Table 4.12. ANNs classification table for TD(-1). 10-7 nodes in the hidden layers.

	Predicted		
Observed	Non-target	Target	%
Non-target	64	14	82.05%
Target	23	55	70.51%
		Overall	76.28%

The ANN model for TD(-1) with two hidden layers (10 in the first hidden layer and 7 in the second) resulted in the best classification rate for the data which generalised well on the test sample. In line with the MD, the network performed better with the non-target sample than the target group. In contrast to the LR model, the network's overall classification rate is considerably higher though number of targets classified correctly is almost similar. It classified 55 targets successfully in comparison with 54 in LR model. On the other hand it classified substantially higher number of non-targets (64) correctly in contrast to LR (45).

Table 4.13. ANNs classification table for IRD. 9-7 nodes in the hidden layers.

	Predicted		
Observed	Non-target	Target	%
Non-target	62	16	79.48%
Target	25	53	67.94%
		Overall	73.71%

As for the IRD, ANN with two hidden layers (9 nodes in the first hidden layer and 7 nodes in the second) provided the best classification results for this data set. Again the network classified more companies into their correct groups than LR model for the IRD, especially with the non-target companies. Overall classification rate is considerable higher for ANN.

Table 4.14. Logistic Regression Models for TD, two years prior to takeover bid, and VD.

Variables	TD(-2)	VD
RSHF	-0.0104	-0.0746
	(1.765)	(1.366)
TUT	-0.00752	-0.025
	(0.688)	(5.993)**
GEAR	0.000628	0.0181
	(0.041)	(0.0987)
LIQD	-3.2117	-0.7555
	(7.281)***	(0.657)
LOG(SIZE)	0.38	-9.5456
	(1.452)	(4.351)**
PROPFIX	-1.2733	0.8329
	(1.305)	(0.879)
P/E	0.00176	0.0327
	(0.036)	(0.3466)
MBV	-0.0122	-0.2599
	(0.624)	(1.848)
GRDUMMY	-0.1147	-0.1203
	(0.095)	(0.133)
Constant	2.3468	0.2361
	(1.802)	(1.486)
N	156	156
Model chi-square	13.79	17.83**
Likelihood Ratio Index	0.0726	0.1270
Correct classification %		
Target	62.82%	66.67%
Non-Target	57.69%	66.67%
Overall	60.26%	66.67%

Notes:

1 ***Significant at the 1% level, 2 tail test

**Significant at the 5% level, 2 tail test

*Significant at the 10% level, 2 tail test

2 Figures in the parentheses are Wald statistics.

3 The likelihood ratio index is defined as (1-log likelihood at convergence/log likelihood at zero). It is similar to the R^2 statistic of multiple regression and gives an indication of the Logit model's explanatory power.

In model TD(-2), the only statistically significant variable is liquidity at the 1% level. As can be seen the model itself is not statistically significant with a very low likelihood ratio index. This result shows that the model does not provide a statistically significant explanation of a firm's acquisition probability. However, it is interesting to note that during the time period between two years and one year prior to the takeover bid, the target companies experienced some changes that increased their likelihood of becoming takeover targets. This can be observed by the significant difference between TD(-1) and TD(-2) models. In TD(-2), other than liquidity, it is not possible to distinguish target companies from non-targets.

As explained above, the result of the minus two years model in contrast to the minus one year model might mean that a change in these variables from year two to one increased the takeover probabilities of target companies. As a result a new model that measures this change from year two to one is estimated for the TD and given the title Variation Data (VD).

The VD model is statistically significant at the 5% level. TUT and SIZE are the only significant variables at the 5% level. The negative coefficient sign of the TUT variable indicates that companies which experienced lower growth had a higher probability of receiving takeover bids during the period examined. Data shows that target companies experienced significant reduction in their growth. On the other hand, however, target companies experienced a significant reduction in size, though their fixed assets proportional to total assets were increased. Even though PROFIX is not significant the sign of the coefficient is positive. This implies that even though the proportion of fixed assets increased as a proportion of the total assets, this happened at the expense of either an increased level of current liabilities or reduced levels of current assets that lead to a reduction in the net assets. It is possible to conclude in line with the positive sign of GEAR, even though it is not statistically significant, that target companies increased their short term borrowing by bank overdrafts that lead to an increase in their current liabilities.

Below tables present the classification rates of ANN models on the derivation/estimation samples of the TD(-2) and VD data samples.

Table 4.15. ANN classification table for TD(-2). (9-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	51	27	65.38%
Target	30	48	61.53%
		Overall	63.46%

The ANN model with 9 nodes in the first hidden layer and 5 nodes in the second hidden layer provided the best results for TD(-2). In this sample ANN has provided better classification overall compared with LR. However the overall classification rate is the lowest in comparison with the earlier samples.

Table 4.16. ANN classification table for VD. (8-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	59	19	75.64%
Target	22	56	71.79%
		Overall	73.71%

The ANN model for the VD shows that the network classified more non-targets, 59 out of 78, than targets, 56 out of 78. Overall, the network provided a better fit to the estimation sample than LR.

4.5.2 Hold-out Test Results

In this section the results of the above models on the hold-out sample will be presented. As explained earlier, the cut-off point for logistic regression is determined *ex ante* from the estimation sample rather than using an arbitrary point of 0.50 (Palepu, 1986). Therefore, first of all the prediction results of the 0.50 cut-off point are presented in comparison with other predictions of estimated cut-off values in order to observe the differences. However as already mentioned earlier the Neural network software that is used in this study only produces binary outputs as targets and non-targets, and it is not possible to obtain individual scores or probabilities as such so that cut-off point can be manipulated. As a result of this limitation only one set of prediction results of ANNs will be presented. This process of estimating error-minimisation by estimated acquisition probabilities for targets and non-targets for time data is presented in detail below. Estimation of this process for the other models is not dealt with in this chapter but the relevant data is presented in Appendix A.

The predictive results of the TD(-2) model will not be attempted since the model itself is not significant and does not provide a statistically significant explanation of a firm's acquisition probability. However the relevant data for probability distribution of targets and non-targets in TD(-2) is presented in Appendix A.

4.5.2.1 Time Data (TD) Results

Table 4.17. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	79	40	66.38%
Target	11	14	56.00%
		Overall	64.58%

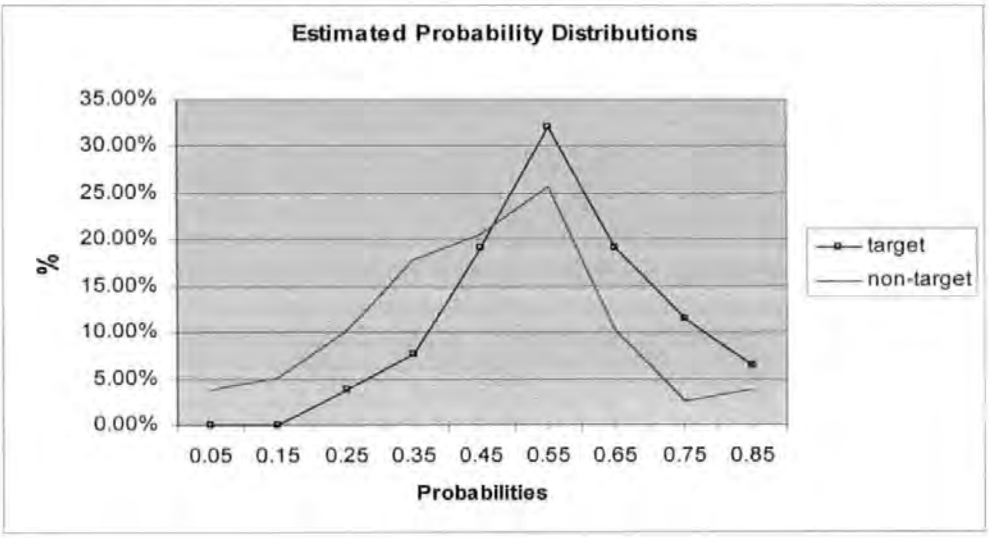
The predictive results of TD are presented in Table 4.17. The overall prediction rate of 64.6% is not greater than proportional chance criterion of 71.3%. The model predicted 93 of 144 companies correctly with an overall predictive rate of 64.58%. The model's classification of non-target companies is higher than the target companies. Unfortunately, only 56% of the target companies are classified correctly compared with 66% of non-targets. The model classified 40 non-target companies as targets along with 14 other companies from the target sample. The total number of companies that are classified as targets are 54 and only 26% of this group are classified correctly as target companies.

In order to estimate the optimum cut-off probability of the model, the marginal probabilities of targets and non-targets in the model development sample are estimated and plotted. The graph is obtained by calculating the frequency distribution of the estimation sample over ten equal intervals of 0.00 to 0.999. The number of targets and non-targets which fall within each interval is obtained and their percentages are computed and plotted on a graph.

Table 4.18. Estimated distribution of targets and non-targets in the TD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	3	3.85%	0	0.00%
0.100-0.199	0.15	4	5.13%	0	0.00%
0.200-0.299	0.25	8	10.26%	3	3.85%
0.300-0.399	0.35	14	17.95%	6	7.69%
0.400-0.499	0.45	16	20.51%	15	19.23%
0.500-0.599	0.55	20	25.64%	25	32.05%
0.600-0.699	0.65	8	10.26%	15	19.23%
0.700-0.799	0.75	2	2.56%	9	11.54%
0.800-0.899	0.85	3	3.85%	5	6.41%
0.900-0.999	0.95	0	0.00	0	0.00

Figure 4.1. Distribution of acquisition probabilities for targets and non-targets



As can be seen from Figure 4.1, the distribution of acquisition probabilities of the two groups intersects at 0.465. It can also be observed from the figure that the separation between the two distributions is not ideal, which means that targets and targets were not separated well by the model estimates. In an ideal situation it is expected that target distribution should mainly concentrate on the right hand side (high acquisition likelihood for targets) while non-target distribution on the left (low acquisition likelihood for non-targets). This cut off value is quite close to the 0.50 and the results based upon this new cut-off probability are presented below.

Table 4.19. LR. Cut-off value of 0.465

	Predicted		
Observed	Non-target	Target	%
Non-target	64	55	53.78%
Target	8	17	68.00%
		Overall	56.25%

The result of Table 4.19 shows that moving the cut-off probability from 0.50 to 0.465 has improved the classification TM model slightly on the target sample but not on the non-target sample. Moreover, the overall classification rate decreased and the overall classification is not enough to exceed the 71.3% chance criterion either. At this cut value, 64 non-target and 17 target companies are classified correctly with an overall classification rate of 56.25%. The result slightly worsened for the group where the targets and non-targets are classified as targets. In this prediction rate a total of 72 companies are identified as targets, increased from 54, and the percentage of real targets in this group is down from 26% to 23.6%.

Table 4.20. ANN. (10-7 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	77	42	64.70%
Target	9	16	64.00%
		Overall	64.58%

The overall prediction result of ANNs is not different from LR. Although ANNs' overall prediction result outperformed LR with a 0.465 cut off value, it did not provide better results when the cut-off point for LR was increased to 0.50. In fact the network has exactly the same overall classification figure of 64.58%. Compared with 0.50 cut-off point, the network predicted more targets correctly but performed lower on the non-target companies. Furthermore the network failed to provide classification rates better than can be expected by chance.

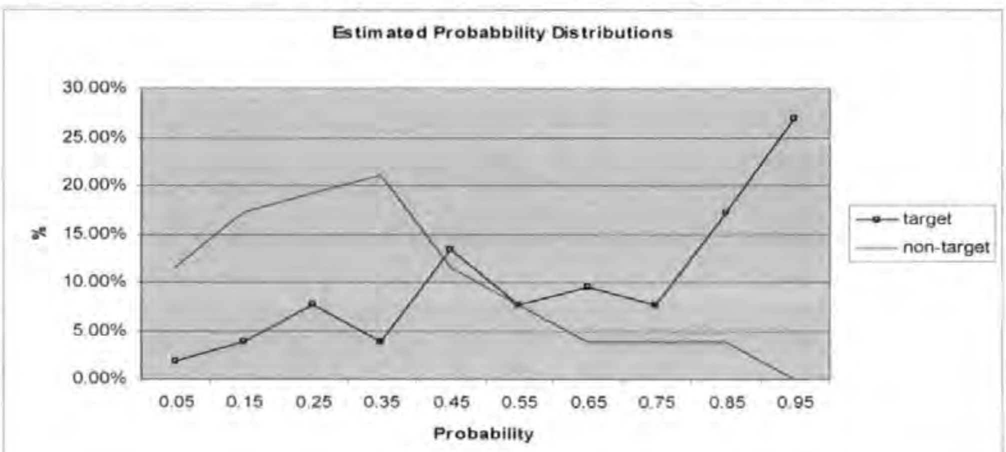
The results of Table 4.19 and Table 4.20 indicate that neither the LR nor the ANNs models provided a classification model with the time data which could be used to develop group profiles. Both models predicted very similar number of targets. Although LR classified 79 out of 119 non-target companies compared with ANNs' 77 non-targets, the difference is not significant.

4.5.2.2 Mixed Data (MD) Results

Table 4.21. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	89	56	61.38%
Target	17	34	66.67%
		Overall	62.76%

Figure 4.2. Distribution of acquisition probabilities for targets and non-targets



The two distributions intersect at 0.44. Figure 4.2 indicates a better separation of distribution of probabilities between targets and non-target companies compared with the TD(-1).

Table 4.22. LR. Cut-off value of 0.44

	Predicted		
Observed	Non-target	Target	%
Non-target	81	64	55.86%
Target	15	36	70.59%
		Overall	59.69%

The overall predictive ability of the LR model has not been improved by choosing the new cut-off probability. At the 0.44 cut off value, the model's prediction rate is lower than the proportional chance probability of 61.5%. On the other hand the overall classification rate at the 0.50 cut off level is higher even though the moving of the cut-off point to 0.44 has improved the accurate classification rate of target companies. However the drop in the correct classification of non-targets companies reduced the overall rate of successful classification. Because the overall classification rate at the 0.50 cut off level is higher than the proportional chance probability of 61.5% *t*-statistic is calculated to see if it is significantly higher than the proportional chance probability. The *t*-statistic for the classification accuracy of LR at the 0.50 cut value is 0.364 and is not statistically significant. The estimation of the cut-off probability of 0.44 is presented in Appendix A. The cut-off probability of 0.44 has classified 81 of 145 non-targets and 36 of 51 targets correctly with an overall classification rate of 59.69%. This is an drop from 62.76% by 0.50 cut value. Again 100 companies are classified as targets at the 0.44 cut-off point and only 36% of them are real targets.

Table 4.23. ANN. (8-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	89	56	61.37%
Target	14	37	72.54%
		Overall	64.28%

As the result Tables of both models indicate none of the models outperformed the other one significantly enough to be called a better modelling technique with MD. At the 0.50 cut value, the LR model classified 89 of 145 non-targets and 34 of 51 targets correctly with an overall classification rate of 62.76%. On the other hand, although ANNs outperformed LR at the 0.44 cut off value, it did not improve the classification rate much for the MD in comparison to 0.50 cut off value. Despite the fact that the overall classification rate is higher than the proportional chance probability of 61.5%, the estimated t-statistic, 0.812, is not statistically significant. The total target companies classified in the target group with LR are 36% compared with 39.8% using ANNs. Although the overall classification of ANNs is slightly higher compared with LR, the overall results are not significant.

The explanatory power of models with MD is higher than the models with TD and the considerable difference in prediction results are as expected. The results of the models with MD are as a result of the formation of the sample split between the estimation and hold-out sample. In terms of percentages, the MD model predicted more target companies correctly but performed slightly worse in predicting non-target companies. As emphasised earlier, MD is used as a benchmark to measure and compare how much the prediction ability of models with TD would be reduced. It shows that even in a time period of one year target characteristics change so that the prediction results are affected.

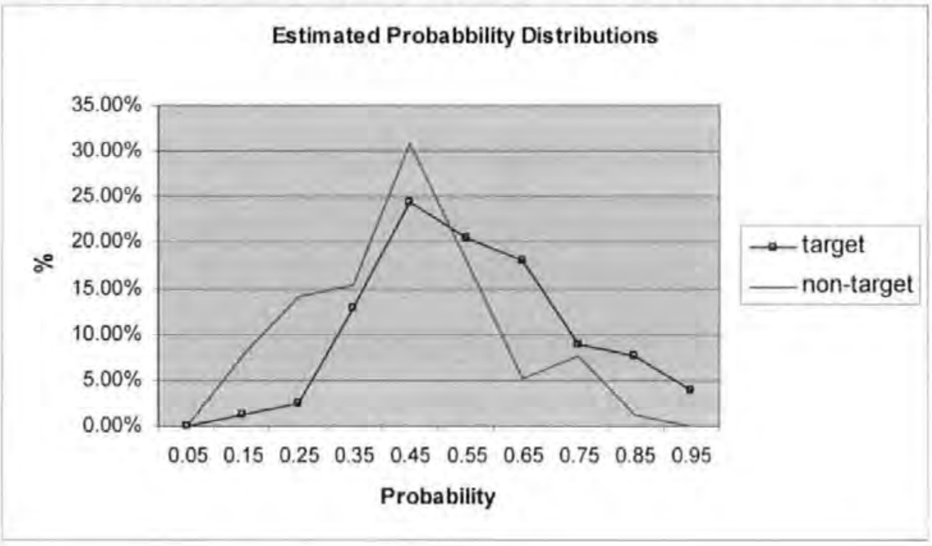
4.5.2.3 Industry Relative Data (IRD) Results

Results of the LR and ANNs models with IRD prediction on the hold-out sample are presented below. In this section, as above, the results of the 0.50. 0.55 cut off points and the result of the ANN will be presented. The estimation of the cut-off probability of 0.55 is presented in Appendix A.

Table 4.24. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	59	60	49.58%
Target	9	16	64.00%
		Overall	52.08%

Figure 4.3. Distribution of acquisition probabilities for targets and non-targets



The Figure 4.3 indicates that the two probability distributions intersect at 0.55. As can also be seen from the figure, the separation of two distributions is not well defined compared to MD.

Table 4.25. LR. Cut-off value of 0.55

	Predicted		
Observed	Non-target	Target	%
Non-target	70	49	58.82%
Target	11	14	56.00%
		Overall	58.33%

Table 4.26. ANN.(9-7 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	70	49	58.82%
Target	6	19	76.00%
		Overall	61.80%

None of the above classifications is statistically significant. Interestingly, the ANN model classified more non-target companies, 70 out of 119, as well as target companies, 19 out of 25, compared with the LR model at the 0.50 cut of point. On the other hand the network classified the same number of non-targets while it performed better on the target companies compared to 0.55 cut of point. The percentage of actual targets classified as target for the LR model is 21% (16/76) at the 0.50 cut-off point, 22.2% (14/63) at the 0.55 cut-off point and 27.9% (19/68) for the ANN model. The difference is not significant. Overall, the IRD model's prediction rate is slightly lower than the TD's prediction rate.

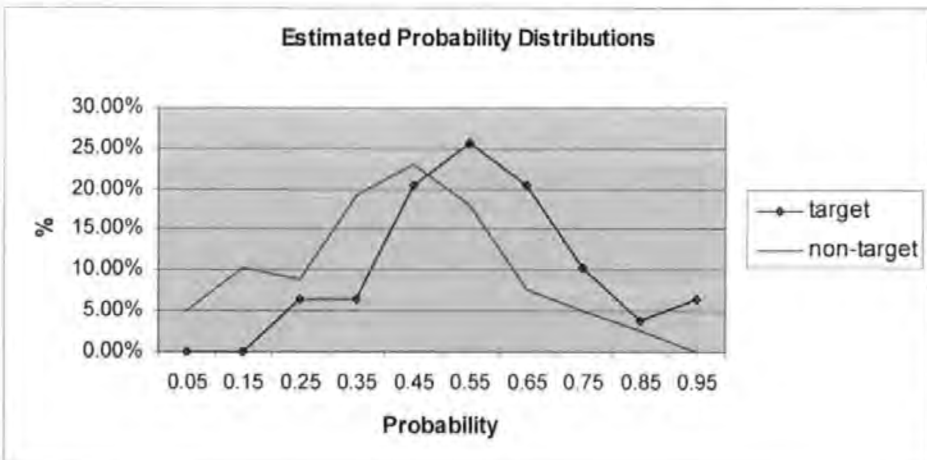
4.5.2.4 Variation Data (VD) Results

Finally, the prediction results of the VD on the hold-out sample is presented below. Although the estimated cut-off point is 0.47, the cut value of 0.50 has performed better classifying companies in the hold-out sample. Therefore, the results of both cut-off points are displayed below. Again the table for the distribution of both targets and non-targets data is presented in Appendix A.

Table 4.27. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	53	66	44.54%
Target	13	12	48.00%
		Overall	45.14%

Figure 4.4. Distribution of acquisition probabilities for targets and non-targets



The above figure shows a better separation of probability distribution compared to TD(-1) and IRD. Non-target distribution peaks up at 0.45 while target population at 0.55. However the two distributions are not well separated.

Table 4.28. LR. Cut value of 0.47

	Predicted		
Observed	Non-target	Target	%
Non-target	46	73	38.66%
Target	12	13	52.00%
		Overall	40.97%

Results show that using the estimated cut-off point did not improve the model's prediction. Although there was hardly any change in the classification of the number of target companies, at the estimated cut-off point the model classified 46 non-target companies, a reduction of 7 from 53 companies.

Table 4.29. ANN.(8-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	64	55	53.78%
Target	12	13	52.00%
		Overall	53.47%

The ANN model has performed better on VD even though the results are not significant. As can be seen in the tables, ANN has classified more non-targets and same number of non-target companies with an overall rate of 53.47%. In terms of the number of real targets correctly classified within all identified targets, LR had 15.4% while ANN had 19.1%. The results of the above three models are quite different from the other models' prediction with their poor prediction results.

In conclusion, VD did not provide a better takeover modelling than the TD model. A simple comparison of two LR models in terms of likelihood ratio index and their prediction results on hold-out samples reveals that the TD model is more robust to takeover modelling than VD. This, of course, is true as far as the employed financial ratios and the time period, where the data is taken, are concerned in the study.

4.6 Discussions and Summary

This chapter has tried to answer the question as to whether it is possible to discriminate between two groups of companies as targets and non-targets on the basis of their financial characteristics. All the estimated models, except TD(-2), were significant at the 1% or 5% levels indicating that they provide a higher explanation of a firm's acquisition probability. However their predictions on the hold-out sample did not produce the desired outcome.

The prediction results of all the models that were applied to takeover predictions show that the TD model has produced the best results compared with the rest of the models. As said in the beginning, the MD model is only used as a benchmark for the TD model in order to be able see the deterioration in the predictive power of the model in mergers and acquisitions where the characteristics which affect the takeover likelihood of target companies change constantly through time. A simple comparison of the two model's statistics reveals that the MD model provided a better explanation of a firm's acquisition probability and prediction results that were higher than can be attributable to the proportional chance probability even though the estimated *t*-statistic for these prediction results were not significant. However, it should be noted here that the higher prediction results of the MD model could be attributable to usage of the same time period in the estimation and hold-out sample as well as the higher significance in the estimation sample which has a likelihood ratio index of 0.29 compared with 0.13 in the TD model.

The overall outcomes of the above models are similar and comparable to the results of the some of the previous models that are discussed below. In these previous studies, however, the success of the models' predictions were measured as the proportion of companies correctly classified and no attempt was made to measure the significance of these overall classification results. On the other hand, it is a simple procedure to compute *t*-statistic for significance for the results of these studies. Some of these previous studies and their results are briefly examined below for comparison. It should be noted here that, as pointed out earlier, those studies which published significant classification rates on one-to-one matched hold-out samples suffer from a bias in their reported error rate. The estimated error rate from the hold-out sample will not be equal to the expected prediction error rate in the population (Palepu, 1986).

Simkowitz and Monroe (1971) reported a significant prediction rate of 64% on the target companies while it was 61% on the non-target sample by applying MDA. They used a hold-out sample of 87 firms which consist of 23 targets and 64 non-target firms. They performed a *t*-test of the proportion of correct classification and found that results were significant at the 5% level.

Stevens (1973) also applied MDA to a sample of 40 (20 target and 20 non-target) firms and reported an overall classification rate of 67.5%. However he did not provide any specific classification percentages over the sub-groups of targets and non-targets. He also did not test the significance of these results.

Belkaoui (1978) constructed his discriminant analysis model on a sample of 50 target and non-target Canadian companies and tested its discriminatory power on a hold-out sample of 11 firms. The results indicate a misclassification rate of 15% or correct classification rate of 85% three years prior to takeover. His results are particularly interesting as he actually found out that the error classification rate for one year prior to takeover was 30%, which means that the model constructed with data three year prior to takeover provided more information to likelihood of firms being takenover.

Wansley and Lane's (1983) linear discriminant analysis model reported a 69.2% classification accuracy on a hold-out sample of 78 firms. Because they used equal matching in their hold-out sample, the reported classification rate is significant (*t*-statistic = 2.54) at the 1% level. These results also suffer from estimated error rates calculated on a matched hold-out sample.

Dietrich and Sorensen (1984) find that their Logit model can predict 90% of the targets and non-targets in the estimation sample and 91% of the targets and 83% of the non-targets in the hold-out sample (*t*-statistic = 3.83). They used a sample of 67 firms (24 merged and 43 non-merged) to built the model and tested it on a hold-out sample of 22 cases (6 targets and 12 non-targets). But the drawing of their estimation and hold-out sample from the same time period makes it very difficult to compare it with the TD model. This result is also biased due to an equal number of targets and non-targets in the hold-out sample.

Barnes (1990) applied multiple discriminant analysis to publicly quoted UK companies and reported that the overall classification ability of the model on the estimation sample was 68.48% and 74.3% on the hold-out sample that was formed by 37 acquired and 37 matched non-acquired companies. However, Barnes (1990) did not attempt to use his model's predictive ability on a subsequent period due to the market crash. This result again under the circumstances of one-to-one matching, is significant (*t*-statistic = 4.18).

In contrast to above-mentioned studies, the prediction results from some other studies, such as Palepu (1986), Sen *et al.* (1995), Barnes (1998), were not so encouraging as they have corrected the issues surrounding the hold-out sample.

Palepu (1986) applied logistic regression technique and estimated the acquisition likelihood on a sample of 419 cases (163 targets and 256 non-target) and tested on a hold-out sample of 1117 firms (30 targets and 1087 non-targets). He found out that the model classified 625 firms to be targets while only 24 were actual targets and 492 companies as non-targets. The prediction results suggest that the model correctly classified 80% of the targets while it only classified 45% of the non-targets correctly. Overall classification rate of 45.6% indicates a very poor prediction and this prediction result is significantly lower than can be expected from a proportional chance probability of 94.8%.

Barnes (1998) reported his results on a sample of UK takeovers. He used a sample of companies from a period of 3 years from 1991 to 1993 and tested it on a hold-out sample comprised of 1994 data. Hold-out sample consist on 16 target companies and 1169 non-target companies. He estimated the cut-off point in a similar method as in Palepu (1986) and experimented different cut-off levels in order to minimise the error classification rate. In most of those predictions with different cut-off points, the models performed badly especially on the target sub-population. None of the models at any cut-off point managed to predict a single target correctly.

The only study that used ANNs in takeover predictions was conducted by Sen *et al.*, (1995) in the USA. Sen *et al* (1995) followed the same methodology of Palepu (1986) in terms of modelling and similar financial ratios. They used 78 targets and 1285 non-targets in their hold-out sample. LR predicted 66.4% of the companies successfully compared with ANNs' prediction of 61.2%. LR classified more in the non-target group (68.6%) than in the target group (29.5%) On the other hand, ANNs performed better with the target group (35.9%) than the non-target group (62.7%). None of the model's prediction rates was significantly higher than proportional chance probability (89.2%) to be considered significant.

Although a relatively smaller hold-out sample is used in this study compared with the one in Sen *et al.*, (1995), the prediction results of both studies are very similar. Even though both techniques performed the same on the target group, LR classified more non-targets. The results of this study in conjunction with Sen *et al.*, (1995) indicate that ANNs neither outperformed LR as a prediction tool nor can be used to predict takeover targets, hence it did not produce any significant prediction results. The advantage of employing ANNs lies in its ability to fit complex non-linear models to the data. On the other hand, the training time is longer when compared with LR, identifying the number of hidden layers and nodes in each layer, and the relative importance of weights can not be known hence the black box criticism. However, employing some of the techniques such as Optimal Brain Damage (OBD) (Cun *et al.*, 1990; Mozer and Smolensky, 1989), training time can be reduced. OBD removes unimportant weights from a network leading to better generalisation, fewer training examples and improved speed of learning.

The results obtained from the five models presented above suggest that it is not possible to predict target companies prior to takeovers by employing LR and ANNs. It should be noted however the results and the shortcomings of ANNs in this study, of course, are limited to the software used in the analysis.

Chapter 5 – Corporate Governance Data Modelling and Results

5.1 Introduction

The internal monitoring of companies is undertaken by their board of directors. Fama, (1980) and Fama and Jensen, (1983) stress that the board of directors is an important internal governance mechanism. Jensen (1986), states that "the internal control mechanism of corporations which operate through the board of directors, generally work well" (p.9). Effective monitoring by the board should result in shareholder wealth maximisation and resolve the agency conflict between management and shareholders. The view that the board is an important internal monitoring mechanism (Fama, 1980; Fama and Jensen, 1983), and the quality of monitoring depends on the composition of the board is important from the perspective of takeovers and takeover prediction. Jensen (1986) argues that the market for corporate control or external takeover market would come into play when internal controls and board level control mechanisms fail or defunct. Jensen's (1986) view suggests or implies that where internal corporate mechanisms are ineffective, the market for corporate takeovers would discipline the management. The findings of Kennedy and Limmack (1996) seem to support this assumption. They have reported that takeovers in the UK act as part of a disciplinary mechanism on inefficient companies, measured on shareholder return, and also found out that the takeovers result in the replacement of inefficient management as they observed a significant increase in CEO turnover two years after the takeover.

In the UK, in order to investigate the problems of financial reporting and effectiveness of auditing, the Cadbury committee was set up. The Cadbury Committee's Report

(1992) outlined the optimum link or interactions between shareholders, board members and auditors and formulated a code based on openness and accountability. The Committee furthermore identified a number of board characteristics that were claimed to represent a good governance practice and this aspect of their conclusion is central to this section of the study. The outlined good governance characteristics included the separation of the roles of chairman and chief executive officer, containing a significant number of independent high-calibre non-executive directors on the board, and the setting up of independent remuneration committees to establish management remuneration packages.

As outlined and articulated by the Committee, where there is such an optimum composition of the board for effective monitoring and hence good financial performance, the failure of this monitoring can be identified by examining the structure of the board. Since the market for corporate control implies that the companies will be targeted for better utilisation of their assets, it can safely be assumed that the companies which become targets for corporate control fail to deliver this optimum board structure and can be identified accordingly (Weisbach, 1988: Shivdasani, 1993).

The corporate governance literature supports the concept that composition of the board of directors is an important element in shareholder' wealth maximisation, especially in transactions where the interests of shareholders and managers may diverge (Rosentain and Wyatt, 1997). Weisbach (1988) reports that the boards that are dominated by outside directors are more likely to replace CEOs in the face of bad performance. The research of Morck, *et al.*, (1988) on how shareholdings by board of directors effect performance show that performance, measured as Tobin's q, is highest at low levels of share ownership by the board. Also they reported that this is the same for he ownership by the firm's top management and for the rest of the board. Their results are supportive of an optimal governance structure. Boyd (1995) looked at the firm performance and separation of chairmanship and chief executive officer and concluded that the separation has a positive effect on firm performance.

The appointment of the Cadbury Committee on corporate governance practices in the UK was as a result of concerns about business failures, creative accounting practices (Smith, 1992), directors' pay especially in the former state-owned corporations, and the highly criticised corporate short-termism (Whittington, 1993). The concerns which lead to the examination of corporate governance practices and the conclusions of the committee (Lumby and Jones, 1999; Forbes and Watson, 1993) are also in support of optimal corporate governance structure for better performance.

Due to the reasons outlined above the qualitative characteristics (non-financial) of companies have received extensive attention from the researchers in finance and accounting particularly in corporate failure predictions. Argenti (1976), for example, argued that financial ratios, because they were simply "symptoms" of business failure, are unable to yield significant insights into the underlying process or "causes" of corporate collapse. Accordingly, Argenti (1976) has developed a dynamic model of business failure which mainly covered the fundamentals of the business and its management structure rather than financial ratios.

Walkling (1985) demonstrated that, in order for a tender offer to be successful, other factors such as managerial resistance, size of the bid premium, use of the solicitation fees, and the initial bidder ownership of targets should be taken into account. Therefore, the success of a takeover attempt is also dependent on factors other than the financial characteristics of firms, and models that include only financial accounting data is seriously misleading (Bartley and Boardman, 1990).

Likewise, in their study of target selection, Gammie and Gammie (1996) stressed the importance of corporate governance variables in a difficult macroeconomic operating environment to firm performance in the UK. In their study, they have included five governance variables. These are: whether the Chief Executive Officer and the Chairman positions are filled by a single individual, the effect of large external shareholders, additional outside directorships held by non-executives, additional outside directorships held by the executive directors, and share options held by the executive directors. They concluded that the influence of governance structures on hostile takeovers is very strong.

Furthermore, Weir (1997) looked at the relationship between the probability of being acquired and firm performance and governance structures in the UK. He made extensive references to the recommendations of Cadbury Committee, especially concerning the board structure and utilised comparable variables to the Gammie and Gammie (1996) study in his analysis. He used seven board characteristics as variables that were; whether the Chief Executive Officer and the Chairman positions are filled by a single individual, proportion of executive directors, sum of the three largest external shareholders, percentage of ordinary shares held by executive directors, mean number of additional outside directorships held by the non-executive directors, mean number of additional outside directorships held by the executive directors, and share options held by the executive directors as a percentage of the total issued ordinary share capital. The study concluded that the internal governance structures of the acquired firms had been ineffective which lead to poor performance.

Findings of these above studies, such as Shivdasani (1993), Gammie and Gammie (1996), and Weir (1997), are discussed in detail in the variables section. The purpose here is to highlight the significance of individual variables/hypothesis employed in this study as well as to link them to the results of these studies.

The objective in this chapter, therefore, is to analyse governance structures, along the recommendations of the Cadbury Committee and the relevant literature, of those companies in the sample. Hence the underlying rationale in the analysis of corporate governance variables in this chapter is that there is a causal relationship between corporate governance or governance structure (measured through board structure) and performance of a company. Therefore it is inferred that poor governance or governance structure will result in poor performance and vice versa. Through this inference, an attempt will be made to develop group profiles of corporate governance characteristics of targets and non-targets for prediction purposes.

5.2 Methodology, Data and Variables

The sample set for the UK companies that were included in the financial ratio analysis is used for the governance data. Therefore same sample sizes in the previous chapter also apply in this section of the analysis. The relevant variables were extracted from Price Waterhouse Corporate Register and annual reports of the related companies. The

Register provides information about company profiles on board composition, the names of executive and non-executive directors, institutional and director shareholdings and the profiles of executive and non-executive directors that list the names of the other companies in which the directors are serving. The earliest issue of the Price Waterhouse Corporate Register dates back to March 1989 and this issue covers 1988 data for the companies. Therefore it is only possible to analyse two years of data for the period of 1990-1996 since only 1988 and 1989 data are available for the companies which were taken over in 1990.

As in the financial variable analysis, the same methodology is followed with two different data sets "Mixed Data" (MD) and "Time Data" (TD) for one and two years prior to takeover. Also, in order to measure the change in the variables from the two to one year period, Variation Data (VD) are estimated from TD. Since there is no industry average presented in any known data sources, analysis of the Industry Relative Data (IRD) for the governance variables is not computed. It can be argued that this average could have been calculated from the sample, the quality of such data and results would have been low since the whole sample only contains approximately 60% of the manufacturing industry.

Although there were no missing observations in corporate governance variables, and they contained few outliers, these outliers were eliminated by winsorizing where the range of the variables is set to a limit of three standard deviations from the mean. The rationale for winsorizing has already been discussed in the previous chapter.

5.2.1 Variables

As in the financial variable analysis of companies, it is important that the selected variables are effective measures of corporate governance. The Cadbury Committee report's conclusions imply that a good governance structure would lead to healthy financial results. An effective governance structure with healthy performance should result in maximisation of shareholders' value and reduce the probability for takeovers (Jensen, 1986). Therefore the variables are selected from the relevant literature and according to The Cadbury Committee's recommendations for a good corporate governance structure.

As stated by Shivdasani (1993), under the corporate governance structure three factors can contribute to the imperfect control of managerial actions. These are, the composition of the board of directors, the structure of equity ownership, and the characteristics of the directors.

The selected variables in this research are complementary to the two other recent takeover studies (Gammie and Gammie, 1996; Weir, 1997) that used corporate governance variables in their analysis of UK public companies and Shivdasani (1993) in the US. The difference in these two UK studies is that their sample included companies quoted on the London Stock Exchange and covered all sectors of the economy.

Some of the limitations of the variables used in this section are that they do not take into account the varying circumstances and different experiences of companies. It can be postulated that different governance structures are optimal for different firms and vary between different industries and sizes. Larger firms, for example, are likely to have more non-executives because of their expertise in monitoring and project evaluation (Shivdasani, 1993). The effect of the industry is controlled by matching companies according to their four digit SIC codes but the size effect still remains. However, the analysis in the previous chapter demonstrated that size is not statistically significant characteristic between target and non-target companies. The Committee on Corporate Governance (1998) stated that the Cadbury Committee report lays down guidelines for corporate governance that are appropriate in most cases and not to be treated as set rules.

The independent variables are:

SepChair: Whether the position of chairman and chief executive are separated. It is identified as 1 if the roles are separated and 0 if not. The expected sign of the coefficient is negative as the separation of the roles is likely to produce efficient running of the board, eliminating or reducing the possibility that the board is dominated by one strong individual. Malette and Fowler (1992), show that the adoption of defensive measures such as position pills is reduced if the roles of CEO and chairman are separated. Rechner and Dalton (1991), reported that the firms which separated these roles had higher financial performance than the ones did not. This is in line with the Cadbury proposal that recommends a separation of these two posts in boards. The Cadbury Committee (1992) recommended that if the internal monitoring mechanisms of the companies to operate effectively, it is important that no individual should held too much power. The committee concluded that separation would diffuse the power concentration in one person and increases the ability of the board to exercise effective control.

The theoretical foundations of separating the roles of CEO and chairman comes from Fama and Jensen (1983b). As put forward by them, an organisation's decision process consists of decision management (initiation and implementation) and decision control (ratification and monitoring). Fama and Jensen (1983a) argue that agency cost in large organisations is reduced by institutional arrangements that separate decision management from decision control. In order to prevent or minimise agency problems initiation and implementation functions of decision making should be separated from the ratification and controlling functions. The main objective of separating the decision making functions of agents from the ratification and monitoring functions is to minimise the agency cost, which inevitably occurs in open public corporations where ownership and management functions are separated. For example, in order to make the Chief-Executive of an company, who is responsible for the initiation and implementation functions of management, more accountable to the board of directors, their duties and responsibilities should be separated from the chairman's and the two posts should be occupied by different people. On the other hand though Brickley *et al.*, (1997) argued that the cost of separating the roles of chairman and CEO is larger than the benefits for most large firms and disagreed with the separation notion.

Shivdasani (1993) reported that the combining the roles of CEO and Chairman reduces the probability of hostile takeover bids while Gammie and Gammie (1996) found out that in both hostile and non-hostile samples, the separation was significant and target companies were more likely to combine the roles. The results of Weir (1997) also support the conclusions that target firms are more likely to combine the roles.

NEX: The proportion of outside directors (non-executive) on the board. The expected sign of this variable is negative. As the proportion of non-executive directors increases in the board, the more likely the efficient monitoring of the executive directors will take place, which will lead to a better financial performance. The Cadbury Committee proposes at least three non-executive directors on the board.

Brinkley and James (1987) reported that outside directors play an important role in the assessment of possible target firms. Moreover Shivdasani (1993) reported a negative relationship between the probability of a hostile takeover and the number of additional outside directorships, even though the numerical representation of outside directors on the board has no significant effect on hostile takeover likelihood. On the other hand, Weisbach (1988) shows that firms with outsider dominated boards are significantly more likely to replace the CEO on the basis of performance. Weisbach's (1988) results give strong support to internal monitoring role of outside directors. The Bank of England's quarterly bulletin (1983 and 1985) emphases the role of non-executive directors as "A suitable non-executive director will generally be able to offer detached and independent advice ... provide additional expertise in specific areas, such as finance. The wider perspective that a non-executive director can bring to bear may be particularly relevant in decisions involving company strategy, or when events of special importance to the company's future, such as mergers and acquisitions, are under discussion" (p.66).

Weir (1997) found the proportion of non-executive directors on board to be a significant at the 10% level on both hostile and non-hostile sample while Shivdasani (1993) did not find it to be significant whether the outside directors were affiliated with the incumbent management or not.

NEXOutDir: The weighted average of outside directorships held by the non-executive directors. The number of outside directorships held by non-executive directors is weighted by the number of non-executive directors on the board. This is estimated as; number of non-executive directorships held in other companies (outside the sample company) by non-executive directors / number of non-executives on the board. The expected sign of this variable is also negative as the proportion of outside directorships held by the non-executive directors increases, the more likely it is that they are valued for their managerial abilities by the market. This is quite an important point since evidence in the US and UK suggests that most of the non-executive directors are in close affiliation with the directors (Forbes and Watson, 1993) and their appointment as non-executives is as a result of this affiliation. Higher value of this variable is an indication of a formal and transparent procedure where appointments of directors are made for their managerial abilities rather than their affiliation with the incumbent management. However, it is also expected that it will be correlated with NEX, as NEX increases NEXOutDir will also increase.

Weir (1997) reported that the mean number of additional outside directorships held by the non-executive directors is significant at the 5% level and Gammie and Gammie (1996) found this variable to be significant at the 5% level for hostile target group while not significant for the non-hostile sample. Shivdasani (1993) also showed that additional outside directorships held by outside directors was significant at the 1% level.

XOutDir: The weighted average of outside directorships held by the executive directors. The number of outside directorships held by executive directors is weighted by the number of executive directors on the board. This is estimated as; number of non-executive directorships held in other companies (outside the sample company) by executive directors / number of executives on the board. The expected sign of this variable is also negative. The same reasons and the logic proposed for theNEXOutDir also apply to this variable.

This variable is important as it will reveal whether executives appointed to the board as a result of their managerial competence and skills or in order to support the existing directors (for their entrenchment purposes). Resenstein & Wyatt (1997) investigated this issue and found that unless the managerial and shareholders' interests are aligned, the market threats an appointment of an insider manager to the board as management entrenchment. Therefore, the higher the proportion of outside directorships held by the executive directors, the better the market values of these executive managers.

Except Shivdasani, (1993), none of the pervious studies mentioned above (Gammie and Gammie, 1996; Weir, 1997) found this variable to be a significant factor between target and non-target companies as well as across hostile and non-hostile samples.

DirShHo: Percentage of the companies' ordinary shares held by the executive directors. Higher shareholdings by the executives will result in a reduction of the agency cost, as this will align the interests of executives with shareholders (Jensen and Meckling, 1976; Shivdasani, 1993). A higher shareholding will have an influence on the likelihood of receiving a takeover bid from two angles. First, as mentioned, it will align the interests of shareholders' with managers (alignment); and second, executives will have an increased saying in the outcome of a possible takeover proposal due to their increased shareholdings (managerial entrenchment). This will also increase the acquisition cost to the acquirer. The expected sign of this variable is negative.

Song and Walkling (1993), showed that managerial ownership is significantly related to the probability of being a target. Their results indicated that target firms have significantly smaller levels of managerial ownership than the non-targets.

Mikkelson and Partch (1997), reported that the target firms have significantly lower insider shareholdings than non-target firms. However the results of Morck*et al.*, (1988) show that the relationship between the shareholdings of executive directors and shareholders' wealth maximisation is not linear. They reported that at high levels of ownership the corporate performance, measured by Tobin's q, falls. Their result may be an indication of managerial entrenchment where the managers gain so much power, at high levels of managerial ownership, that they act to further their own interest rather than that of shareholders.

Weir (1997) showed that lower the proportion of a firm's ordinary shares owned by the executive directors, the higher the probability of becoming a takeover target. Gammie and Gammie (1996) also found that share options held by executive directors are conversely related to the probability of receiving bids. On the other hand, Shivdasani (1993) did not report any significant relationship between equity ownership of management and takeover bids.

BigShHo: Outside shareholdings of major shareholders. This is measured as a total percentage of shareholdings held by outside interests other than those held by board of directors and since the majority of the big shareholders are institutional shareholders it can normally be assumed that the results can be applied to institutional shareholders as well. In theory large shareholders are considered to be effective monitors as they have a vested interest in minimising asymmetry of information and act in accordance with their own interest (Jarrell and Poulson, 1987). Nevertheless, Ambrose and Megginson (1992), Davis and Stout (1992), and Weir (1997) found that outside shareholder or institutional shareholders had no effect on the probability of becoming a takeover target. On the other hand though, Shivdasani (1993) found that the large external shareholdings significantly increased the probability of receiving a hostile bid. Even though the previous research contains conflicting results, the expected sign for this variable is negative because large shareholders have an incentive to monitor the management and any dissatisfaction will result in selling their shares. Some research (Shleifer and Vishny, 1986) supports the notion that large blocks of shareholders monitor managers more effectively due to their increased interests. On the other hand, recently institutional investors have been criticised in the financial press for not doing enough monitoring. Pound (1988), for example, highlighted the fact that institutional shareholders monitoring role might be minimal if they choose to sell their shareholding rather than a positive engagement with the management to improve the its performance. Obviously this variable also contains the shareholdings (toehold) of bidders in the target sub-sample. However this cannot be utilised as a separate variable as this information does not exist for the matched non-target sub-sample.

Table 5.1. Corporate Governance Variables and Expected Signs

Hypothesis	Exp. Sign
Whether the positions of chairman and chief executive are separated (SepChair)	-
The proportion of outside directors on the board (NEX)	-
The weighted average of outside directorships held by the non-executive directors (NEXOutDir)	-
The weighted average of outside directorships held by the executive directors (XOutDir)	-
Percentage of the companies' ordinary shares held by the executive directors (DirShHo)	-
Outside shareholding of major shareholders (BigShHo)	-

As in chapter four, the aim in the first part of this chapter is to identify and model those corporate governance differences between two groups of companies by those above variables as targets and non-targets by employing Logistic Regression (LR) and Artificial Neural Networks (ANNs). In the second part the results of these models will be compared.

5.3 Univariate Data Analysis Results

Tables 5.2, 5.3 and 5.4 provide the summary of descriptive statistics for the MD, TD, and TD(-2). As in the previous chapter along with mean and *t*-values, median and Wilcoxon non-parametric test for median statistics are also provided. In addition to univariate analysis, correlation matrixes are supplied in Appendix C.

Table 5.2. t-statistics: Mixed data (MD) statistics. One year before the bid.

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
SepChair	0.5192	0.7692	0.000	1.000	-3.05***	-2.64***
NEX%	0.3596	0.4057	0.400	0.400	-1.26	-0.61
NEXOutDir	1.3567	1.4154	1.000	1.000	-0.24	-0.30
XOutDir	0.2327	0.7179	0.000	0.105	-2.62**	-1.45
DirShHo	9.9306	14.7717	4.780	9.110	-1.67*	-1.34
BigShHo	29.9269	28.0317	26.81	32.04	0.51	0.43

Notes:
(1) ***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
(2) Observations (52 Targets and 52 Non-targets)

Table 5.2 shows that SepChair, at the 1% level, XOutDir, at the 5% level, and DirShHo, at the 10% level, are significant variables in MD. However, according to the median test of Wilcoxon Z, only variable that is statistically significant at the 1% level is SepChair. The results indicate that target companies combine the roles of chief executive director and chairman. Also, *t*-statistic indicate that executive directors of target companies hold less outside directorships in other companies compared with the matched sample. Furthermore these directors are less aligned to the interest of their shareholder as they hold significantly less shares in companies that they serve. The only variable that has a higher mean value for target companies is the percentage shareholding of big investors, however the difference is significant neither in *t*-test nor in Wilcoxon Z test.

Table 5.3. t-statistics: Time data (TD-1) statistics. One year before the bid.

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
SepChair	0.4872	0.7179	0.000	1.000	-3.27***	-2.93***
NEX%	0.3320	0.3680	0.300	0.300	-1.39	-1.04
NEXOutDir	1.1115	1.3368	0.790	1.000	-0.97	-0.66
XOutDir	0.4279	0.4284	0.105	0.000	-0.05	0.45
DirShHo	12.7790	13.4908	5.0300	6.6200	-0.29	-0.26
BigShHo	30.0967	23.9979	28.5500	19.9100	2.04**	1.95*

Notes:
(1) ***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
(2) Observations (78 Targets and 78 Non-targets)

In TD, SepChair and BigShHo are the only significant variables at the 1% and 5% levels while BigShHo is significant at the 10% level with Wilcoxon test. XOutDir and DirShHo are no longer significant in this data set. However, target companies still have lower mean scores in all variables except in BigShHo. A significant difference in SepChair between targets and non-targets is again an indication that in the target companies the roles of chairman and CEO are combined indicating that the board is dominated by a single individual, in line with lower non-executives and their lower outside shareholding. On the other hand the significance of BigShHo indicates that target companies have higher percentage of big investors compared with non-target ones.

Table 5.4. t-statistics: Time data (TD-2) statistics. Two years before the bid.

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
SepChair	0.5769	0.6795	1.000	1.000	-1.30	-1.31
NEX%	0.3576	0.3628	0.300	0.400	-0.16	0.46
NEXOutDir	1.2744	1.2437	1.000	0.775	-0.66	0.25
XOutDir	0.3944	0.4388	0.210	0.000	-0.45	0.01
DirShHo	16.2494	14.4267	6.5700	9.1200	0.68	0.10
BigShHo	23.7206	21.9296	24.7500	19.1950	0.65	0.65

Notes:
(1) ***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
(2) Observations (78 Targets and 78 Non-targets)

Table 5.4 shows that none of the second year variables is statistically significant. This is an interesting result since SepChair is significant at the one percent level one year prior to takeover. This shows that during the time period from two years to one year prior to takeover a change in SepChair took place to make this variable statistically significant. The mean score for targets dropped from 0.576 to 0.487 compared to an increase of 0.0385 from 0.679 to 0.718 for the non-targets. Target companies still have a lower mean value for SepChair two years prior to takeover in relation to non-targets but this difference is not statistically significant.

Table 5.5. t-statistics: Variation data (VD) statistics.

Ratios	Target Mean	Matched Mean	Target Median	Matched Median	t-values	Wilcoxon Z
SepChair	-0.0897	0.0385	0.000	0.000	-1.59	-1.52
NEX%	0.1230	0.1043	0.000	0.000	-0.96	0.21
NEXOutDir	-0.1629	0.0932	0.000	0.000	-1.37	-1.07
XOutDir	0.0332	-0.0103	0.000	0.000	-0.21	0.24
DirShHo	-3.4704	-0.9359	0.000	-0.010	-1.40	-0.19
BigShHo	6.3760	2.0683	3.800	0.000	1.65	1.67*

Notes:
(1) ***Significant at the 1% level, 2 tail test, **5% level, 2 tail test, *10% level, 2 tail test
(2) Observations (78 Targets and 78 Non-targets)

None of the variables are statistically significant in the variation data accept the increase in the shareholdings of big investors in the target companies (BigShHo) with Wilcoxon Z test. The target companies experienced a reduction in the mean score of SepChair, NEXOutDir, and DirShHo. This implies that two to one year prior to takeover, separated roles of chairman and CEO are combined in target companies while the process was reversed in non-target companies. Although there was an increase in the proportion of non-executives in target companies compared to a lower increase in non-targets, there was a decrease in the outside shareholding of these non-executive directors while a slight increase can be observed in non-target companies. This increase in non-executive directors may be to counter the perceived ill effects of combining the roles of chairman and CEO in the eyes of investors and financial markets alike. However, even though the differences in changes are not significant, the results imply that quality of non-executive as well as executive directors declined in target companies during this period. It is obvious from the numbers above that an increase in the proportion of non-executives in target companies did not bring the quality that was required. It is also interesting that there was an increase in the shareholding of big investors especially in target companies. This increase is statistically significant at the 10% level with the median test of Wilcoxon Z. This might be as a result of a possible takeover rumour or expectation of such a takeover bid.

5.4 Multivariate Models and Results

The LR estimation of the three models is presented below. Table 5.6 shows the MD as well as one year TD(-1) and two years TD(-2) prior to the takeover bid. However a meaningful estimate of Variation Data (VD) could not be estimated. Logistic Regression estimate come up with a warning that 'There is a complete separation of data points. The maximum likelihood estimate does not exist. Validity of the model fit is questionable'. Nonetheless the logistic procedure continued in spite of the warning but as the results are questionable they will not be reported here. However the estimates of VD can be seen in Appendix E. As in the previous chapter, classification tables and percentages of ANNs on the estimation sample are presented under LR models.

The same procedure of obtaining the network structure is followed in this section. The estimation sample is divided into two as training and validation and tested on the hold-out sample. Many trials of single and multiple hidden layers with different nodes in each layer were conducted. Similarly, two hidden layer networks provided better generalisation on hold-out samples. The sigmoid activation function with 0.1 learning rate and 0.8 momentum value is used in the training.

5.4.1 Estimated Models

Table 5.6. Logistic Regression Models for MD, TD. One and two year prior to takeover bid.

Variables	MD	TD(-1)	TD (-2)
SepChair	-1.2198	-0.9473	-0.4644
	(6.595)**	(7.000)***	(1.726)
NEX%	-0.0825	0.0438	-0.0707
	(0.306)	(0.110)	(0.400)
NEXOutDir	-0.0427	-0.0595	-0.0311
	(0.536)	(1.437)	(0.426)
XoutDir	-0.2912	-0.00242	-0.0192
	(3.998)**	(0.001)	(0.0723)
DirShHo	-0.0462	0.00107	-0.00549
	(5.783)**	(0.008)	(0.2842)
BigShHo	-0.00898	0.0185	0.00761
	(0.567)	(3.543)*	(0.627)
Constant	2.3875	0.1725	0.00641
	(8.228)***	(0.105)	(0.0002)
N	104	156	156
Model chi-square	16.08**	14.52**	3.32
Likelihood Ratio Index	0.136	0.070	0.015
Correct classification %			
Target	69.23%	58.97%	50.00%
Non-Target	69.23%	64.10%	58.97%
Overall	69.23%	61.54%	54.49%

Notes:

1 ***Significant at the 1% level, 2 tail test

**Significant at the 5% level, 2 tail test

*Significant at the 10% level, 2 tail test

2 Figures in the parentheses are Wald statistics.

3 The likelihood ratio index is defined as (1-log likelihood at convergence/log likelihood at zero). It is similar to the R^2 statistics of multiple regression and gives an indication of the Logit model's explanatory power.

Apart from the TD(-2), the two other models are statistically significant at the 5% level. As in the financial modeling in the previous chapter, the MD has a higher explanation of acquisition probability with 13.6% likelihood ratio index. In contrast to financial ratio modelling, the above models are less significant in terms of their chi-square and likelihood ratio index hence provide a lesser explanation of a firm's acquisition probability.

In MD, SepChair, XoutDir and DirShHo are statistically significant at the 5% level. This indicates that combining the roles of chairman and CEO, lower numbers of directorships held by executive directors, and lower shareholdings of company directors resulted an increase in probability of receiving a takeover bid. Therefore the target companies in the estimation sample combined the roles of chairman and CEO leading to one man dominated board. Target companies also had executive directors whose managerial abilities were not highly rated by the market and whose appointment was seen by the market to be a result of their affiliation with the company directors or management entrenchment. Statistically significant difference of DirShHo means that, in comparison with the non-target sample target management interest were not aligned to their shareholder and this in turn lead to poor performance and eventual takeover bid. The results are in line with the findings of Song and Walkling (1993), Mikkelson and Partch (1997), and Weir (1997).

In TD(-1), however, XOutDir or DirShHo are no longer significant. In fact the only statistically significant variable other than SepChair is BigShHo. Again the target companies in the sample combined the roles of chairman and CEO and concentrated two much power in the hands of one person. This finding is in contrast to the reported findings of Brickley *et al.*, (1997) that the cost of separating the roles of chairman and CEO is larger than the benefits for most large firms and is in strong support of one of the recommendations of the Cadbury proposals that refer to the roles of chairman and CEOs. The positive sign and significance of BigShHo means that target companies had higher percentage of shareholdings by big investors. However the inference that can be drawn from this result is not what is expected as it is generally assumed that large shareholders are effective monitors with their increased interest in companies that they invest. This result is in line with Shivdasani (1993) as he found out that large external shareholders significantly increased the probability of receiving a bid. The models explanatory power is lower than MD with likelihood ratio index of 7.0%.

The model formed by using TD(-2) is not statistically significant and does not explain a firm's acquisition probability. Even the overall classification accuracy of the model on estimation sample is not statistically significant. Although a 54.49% classification rate is higher than would be expected from chance probability of 50% for an equal sample size, a *t*-statistic of 1.1 is not statistically significant. Hence this model does not provide a statistical explanation of acquisition probability; it will not be used to test on a hold-out sample for prediction purposes. The insignificant result of the second year model is somehow not surprising from the point of view that the market would have reacted earlier, if the differences between target and non-target were significant two years prior to the bid. This assumption, of course, depends upon the efficiency of the market to operate.

As can be seen in TD(-1), for example, the coefficient sign of NEX is positive, though not statistically significant, indicating that the higher proportion of non-executive shareholding will increase the likelihood of a company to become a target. Although the coefficient sign of the same variable in TD(-2) is negative, again it is not statistically significant. Therefore, it is imperative to see the changes in the governance variables. One of the limitations of using cross sectional analysis of single year governance data is that it becomes too static. As pointed out by Hermalin and Weisbach (1991), cross sectional regression of performance on board composition will be biased because of changes in board composition resulting merely from past performance. However the results of two TD models suggest two possible reasons for the coefficient sign change in NEX%. First, this is as a result of information asymmetry between managers and investors. It may be that the managers of the target companies in anticipating the oncoming reduction in their company's performance attempt to change their board structure in order to be seen to be doing something positive in the face of an anticipated underperformance. Second, as mentioned before, in order to eliminate the negative perception of combining the roles of chairman and CEO.

As a result of the difference between two models of TD(-1) and TD(-2), an attempt was made to form a fourth model that measures the change in the variables from TD(-2) to TD(-1). However as mentioned above Variation model could not be estimated. Logistic regression analysis could not estimate the coefficients due to complete separation of data points. The parameter estimates and the validity of the model were in question. The model, along with the warning message, is presented in Appendix E.

Again the tables below present the classification rates of ANN models on the derivation/estimation samples. As it was mentioned in the previous chapter, the classification rates provide mathematical fit of the network models on the estimation samples. They are useful in comparing to LR classification on the estimation samples.

Table 5.7. ANNs' classification table for MD. 8-6 nodes in the hidden layers

	Predicted		
Observed	Non-target	Target	%
Non-target	40	12	76.92%
Target	10	42	80.76%
		Overall	78.84%

ANNs with 8 nodes in the first hidden layer and 6 nodes in the second hidden layer provided the optimum results for MD. As can be seen the overall classification of the network is 78.8% and more target companies are correctly classified than non-target companies. Compared with LR, the networks' fit to the data is higher.

Table 5.8. ANNs' classification table for TD(-1).6-5 nodes in the hidden layers.

	Predicted		
Observed	Non-target	Target	%
Non-target	58	20	74.35%
Target	24	54	69.23%
		Overall	71.79%

The two hidden layer network with 6 nodes in the first layer and 5 nodes in the second hidden layer provided the best model on TD. Compared with the MD data, the network classified more non-target companies correctly than targets with an overall classification rate of 71%.

Table 5.9. ANNs' classification table for TD(-2).6-5 nodes in the hidden layers.

	Predicted		
Observed	Non-target	Target	%
Non-target	49	29	62.82%
Target	23	55	70.51%
		Overall	66.66%

The networks provided a better classification than LR on the TD(-2). As mentioned before ANNs are able to provide a mathematical fit to complex non-linear data even with a single layer hidden structure. But this fit, usually, does not provide good results on the hold-out sample, because the network loses its generalisation ability. As can be seen from the classification result of the estimation sample the network trained and classified better on the target group than the non-target group.

5.4.2 Hold-out Test Results

In this section the prediction results of LR and ANNs models on hold-out samples of above data sets will be presented. As in the previous chapter, the classification results of the estimated cut-off points along with the 0.50 cut value will be shown and the process of estimating the optimum cut-off probability will be presented in Appendix B.

Since TD(-2) does not provide a significant explanation of a firm's acquisition probability, this model will not be used for prediction purposes. However the relevant data for the distribution of two groups in TD(-2) is presented in Appendix B.

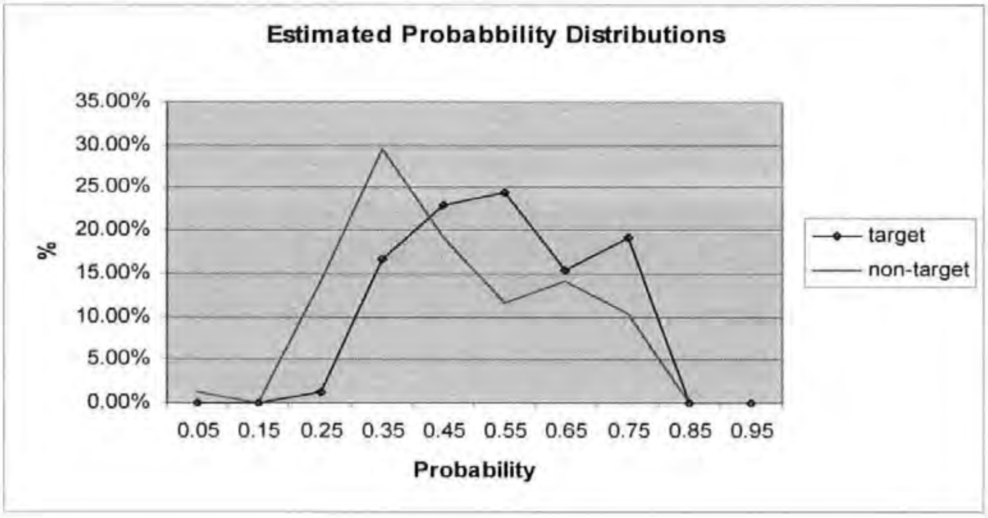
Also the prediction results of the models will be compared with the proportional chance criterion and their significance, in the event that overall prediction exceeds the proportional chance probability, and will be measured with t-statistic in a similar way to that explained in the previous chapter. The proportional chance probability is 71.3% for TD and 61.5% for MD. In the event that the overall prediction percentages of the models exceed these chance probabilities a t-test for significance will be applied.

5.4.2.1 Time Data (TD(-1)) Results

Table 5.10. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	79	40	66.38%
Target	17	8	32.00%
		Overall	60.41%

Figure 5.1 Distribution of acquisition probabilities of targets and non-targets



As can be seen the probability distribution of targets and non-targets in the estimation sample intersect at 0.429. The classification result of this cut-off point is presented below.

Table 5.11. LR. Cut-off value of 0.429

	Predicted		
Observed	Non-target	Target	%
Non-target	53	66	44.53%
Target	11	14	56.00%
		Overall	46.52%

At the 0.429 cut value the overall classification rate has actually dropped compared with the 0.5 cut-off probability. Although 6 more companies are classified correctly in the target group compared with 0.5 cut-off probability, correct classification of non-target companies dropped dramatically to 44.5 percent from 66 percent. The distribution of targets and non-targets in Figure 5.1 suggest that even though the high percentages of non-target companies in the estimation sample are below the estimated probability of 0.42, the classification of this cut-off point is somehow poor on the hold-out sample.

Table 5.12. ANNs. (7-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	84	35	70.58%
Target	14	11	44.00%
		Overall	65.97%

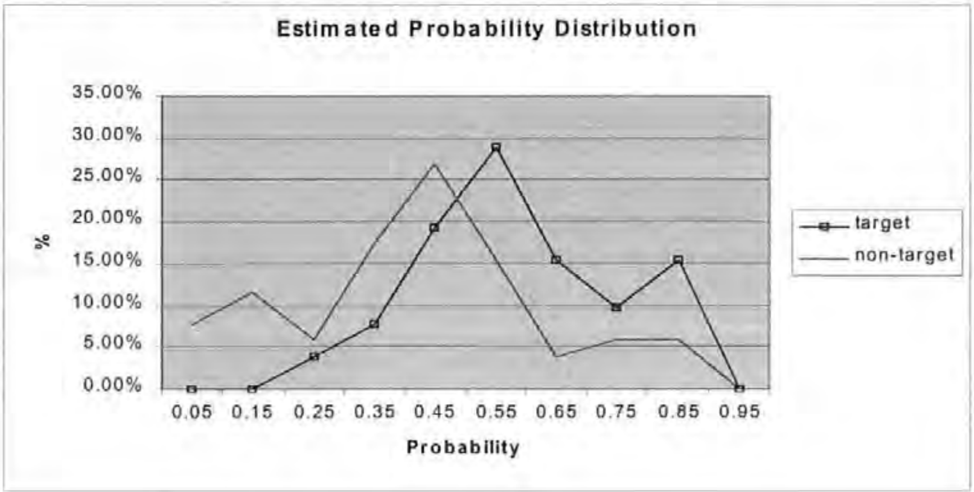
ANNs performed better on both the target and non-target group compared with LR. However, the overall classification of the network is not statistically significant either in terms of proportional chance probability. The correct classification of 23.9% (11/46) actual targets within the target group is also bigger compared to 17.5% (14/80) at the 0.429 cut-off probability.

5.4.2.2 Mixed Data (MD) Results

Table 5.13. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	87	58	60.00%
Target	25	26	50.98%
		Overall	57.65%

Figure 5.2. Distribution of acquisition probabilities of targets and non-targets.



On this occasion the distributions of the two groups intersect at 0.48 which is very close to 0.50 cut-off point. As can be seen in Figure 5.2, a high percentage of target companies are concentrated at the probabilities of 0.55 and 0.85. On the other hand, the distribution of non-target companies is skewed to the left peaking at 0.45. The prediction results of 0.48 intersect point is presented below.

Table 5.14. LR. Cut-off value of 0.48

	Predicted		
Observed	Non-target	Target	%
Non-target	85	60	58.62%
Target	22	29	56.86%
		Overall	58.16%

The classification point of 0.48 provided a slightly higher prediction rate than 0.50 cut-off point. At the 0.48 cut value the prediction rate of the model on the target companies increased by 3 while it classified less on the non-target group. The close classification results are not surprising considering that the estimated cut-off probability of 0.48 is very close to 0.50 cut-off point. None of the overall prediction results are above the proportional chance probability of 61.5%. Therefore it can not be concluded that by using MD data significant classification of target and non-target companies can be obtained.

Table 5.15. ANNs. (8-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	93	52	64.13%
Target	23	28	54.90%
		Overall	61.73%

The network's overall classification rate of 61.7% is higher than above two LR classification rates. The network also classified slightly more in the target group, compared with 0.50 cut-off value but less on the 0.48 cut-off level, and 4 percentage points higher in the non-target group. Although the network classification rate is higher than the proportional chance probability, the difference is not statistically significant. Also, the network's classification rate of real targets within the target group is 35% (28/80) compared to 32.5% (29/89) in LR at the 0.48 cut-off point. The difference is again not significant.

5.4. Discussions and Summary

In this chapter, the analysis is conducted to find out whether it is possible to discriminate target companies from non-targets by employing corporate governance characteristics. The estimated models, except TD(-2), were significant at the 5 per cent level indicating that they provide a higher explanation of a firm's acquisition probability. In MD it was found out that separation of chairmanship from CEO, outside shareholding of executive directors, and directors shareholdings in their companies were statistically significant. The negative estimation sign of these significant variables indicate that target companies combined the roles of chairman and CEO, included executive directors in their boards whose managerial skills were significantly underrated by the labour market, and did not align the interests of managers with their shareholders. Furthermore, the negative coefficient signs of other variables confirm the outlined hypotheses and their relationship with the outcome variable in the model. On the other hand, in TD(-1), outside executive directorships and shareholdings of directors in their companies no longer statistically significant. However shareholdings of outsiders in these companies became significant with a positive coefficient sign indicating that target companies had higher big shareholders compared with the non-target ones. Moreover in TD(-1) the estimation signs of percentage of non-executive directors in the board and shareholdings of directors are positive. The result points out, that contrary to the outlined hypotheses, target companies attempted to align the interest of their managers with their shareholders and had a higher percentage of non-executive directors in their board. However none of these variables are statistically significant to merit any comprehensive conclusions in terms of takeover likelihood modelling. As for

TD(-2) though, the model itself is not significant and does not provide any meaningful analysis of a firm's acquisition probability.

As in the previous chapter, none of the models presented statistically significant classification results on the hold-out samples to be considered in the takeover prediction studies. Although one of the ANNs' models provided overall classification rates slightly above the proportional chance probability, these prediction rates were not statistically significant. Therefore it cannot be ruled out that the presented prediction rates could be achieved by chance.

None of the LR models produced prediction results above the proportional chance criterion. On the other hand the prediction rate (61.7%) of ANNs on MD was slightly above the proportional chance criterion. However this overall prediction result was not statistically significant. As pointed out earlier MD is only used as a benchmark for the TD. As in the financial ratio modelling, MD provided a higher likelihood ratio index compared with TD(-1). However, like in the financial ratio modelling, the MD's prediction results were not statistically significant either.

As in the previous chapter with the financial ratios, ANNs did not perform better in predicting target companies compared with LR modelling by using governance variables. Even though the overall classification of the network on both data sets were higher than for LR classification rates, none of the predictions were statistically significant. ANN classified 66% of the companies into their correct groups with TD(-1), which is lower than the proportional chance. LR, on the other hand, managed to classify

60% of the companies correctly. These prediction results are below the proportional chance probability and can be attributable to mere chance.

The classification results obtained from the two models presented above suggest that it is not possible to predict or classify possible target companies beforehand with significant results by employing Artificial Neural Networks or Logistic Regression. However it should be stressed that the results are limited to the corporate governance variables that are employed in this study.

Chapter 6 – Combined Models (Financial and Corporate Governance) and Results

6.1 Introduction

In this chapter, the data used in previous chapters will be combined for modelling and applied to takeover predictions. The combined models below will include financial ratios (financial variables) and corporate governance variables (non-financial) for the companies in the sample set.

In chapter four, takeover prediction is modelled by using financial ratios to discriminate between the target and non-target companies. The rationale behind using financial ratios as a discriminatory tool was explained in the related chapter. But the basic underlying assumption was that if the target companies underperform compared with their counterparts, their underlying financial indicators (ratios) could be utilised to form group profiles that can be used for discriminatory purposes.

Also stated in chapter five is that the reason for using corporate governance variables in takeover prediction modelling was that there is a causal relationship between corporate governance structure and performance. This assumption implied that there is an optimal corporate governance structure and deviations from that would result in poor financial performance. Hence, in an environment where managers compete for better utilisation of assets then would act upon this and acquire companies that deviate from the optimal governance structure. In this respect it is assumed that the target companies' internal governance structures are organised ineffectively.

Argenti (1976) argued that financial ratios were simply "symptoms" of business failure, and as a result suggested that they were the fundamentals of the business and management structures to be considered for the causes of corporate collapse. Therefore, it can be postulated that financial ratios as measures of corporate performance were simply a reflection upon the corporate governance structure. Therefore, the very reason for target companies to be taken over was that they did not establish proper internal monitoring mechanisms through their corporate governance structures which lead to the external monitoring mechanism (market for corporate control) to come into effect.

Even though some studies (Shivdasani, 1993; Gammie and Gammie, 1996; Weir, 1997) combined these two groups of characteristics of companies to a limited extend for better profiling of target company characteristics, so far there is no example of this type of extensive combination and analysis in takeover prediction studies. Sihivdasani (1993) used financial characteristics (Leverage, Market to book ratio, and Growth rate) as control variables for the board characteristics. On the other hand, Gammie and Gammie (1996) and Weir (1997) only supplemented profit figure to their corporate governance characteristics.

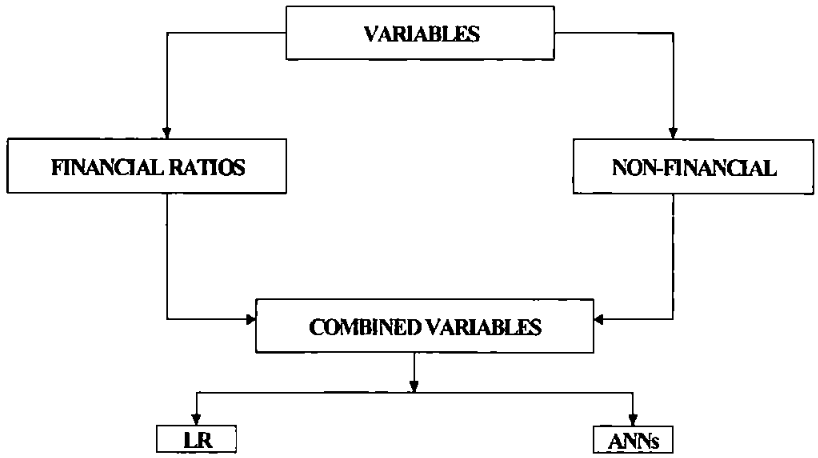
However as the relation between the two set of variables described above, financial ratios as measures of corporate performance were simply a reflection upon the corporate governance structure. This implies a time lag and an interaction between corporate governance characteristics and financial variables. In this section of the analysis, this aspect of feedback from governance change to financial impact or interaction between the two sets of variables will be exploited from takeover predictions modelling perspective. It is expected that this interaction as well as the inclusion of more company characteristics of these companies will increase the estimation power of the models and will lead into better prediction rates. Furthermore, in contrast to the previous chapters, stepwise elimination of combined models will be carried out to see important determinants or characteristics that separate target companies from non-target ones. This analysis will enable us to observe the significant differences between targets and non-targets not only in one aspect of their characteristics such as financial ratios but also in general.

As a result in this section, in order to improve the predictive ability of the models, financial and non-financial variables of the companies will be combined. Similar type of applications or examples of this type of analysis can be found in corporate failure literature (Daily and Dalton, 1994).

6.2 Methodology

The same methodology that was explained in the previous chapters (chapter 4 and 5) is applied in this section. Diagram 6.1 gives a general description of the methodology in this section. As can be seen in diagram 6.1, financial variabks (ratios) and non-financial variables (corporate governance) are combined to see whether the predictive abilities of the models could be improved by taking advantage of the lag effect and interaction between two sets of variables.

Diagram 6.1. Schematic description of the steps involved in the combined methodology.



However, the modelling is conducted only on time and mixed data (TD and MD). The reasoning for this is explained below. In order to identify the data sets used in this chapter time data is called total time data (TTD) while mixed data is called total mixed data (TMD). Also, a stepwise regression analysis is performed. Among the available variable selection methods in SAS, the stepwise elimination was selected.

There are five variable selection methods that are available in SAS. These methods are forward selection, backward elimination, stepwise selection, and score for best subsets selection. The stepwise option is similar to the forward selection option except that variables already in the model do not necessarily remain. Variables are entered into and removed from the model in such a way that each forward selection step can be followed by one or more backward elimination steps. Hence the stepwise selection combines the forward and backward selections methods to add variables to the model or remove variables from the model as they meet or fail to meet specified significance levels, respectively. The stepwise selection process terminates if no further variable can be added to the model or if the variable just entered into the model is the only variable removed in the subsequent backward elimination (SAS System, 1995).

Consequently, the stepwise regression based on the stepwise elimination method with 0.05 statistical significance for inclusion in the model is performed to TMD and TTD. These models are called "reduced time data" (RTD) and "reduced mixed data" (RMD).

The same procedure of network training is followed in this section. The estimation data were separated into training and test samples. A number of different trials were carried out for hidden layers and nodes for these layers to determine the optimum network structure. The sigmoid activation function with a 0.1 learning rate and a 0.8 momentum value is used in the training.

6.3 Multivariate Models and Results.

The LR models for TMD, RMD, TTD and RTD one year prior to takeover are presented in Table 6.1 below. Also, the classification tables of ANNs on the estimation data are shown below in Table 6.2 to 6.5. Correlation analysis for these variables is supplemented in Appendix C.

The reason why only these two data samples (TMD and TTD) are considered for the combined models is simple. Models constructed by using TD(-2) data samples were not statistically significant in both cases of financial and non-financial analysis therefore it is not considered in the combined sample. IRD that was used in the financial analysis was not used in the non-financial section so it is not possible to apply it in this section. The reason for not applying variation data (VD), for example, in this analysis is that it could not have been possible to construct this model in the non-financial characteristics modelling. Therefore VD and TD(-2) are excluded from the analysis in this chapter.

6.3.1 Estimated Models

Table 6.1. Logistic Regression Models for TMD, RMD, TTD and RTD.

Variables	TMD	RMD	TTD	RTD
RSHF	-0.057	-0.0631	-0.0102	
	(6.9801)***	(11.875)***	(1.872)	
TUT	-0.0506		-0.0228	-0.029
	(6.0944)**		(4.418)**	(8.262)***
GEAR	0.0097		0.00267	
	(3.5739)*		(0.468)	
LIQD	-2.1281		-2.3957	-2.3399
	(2.126)		(5.241)**	(7.504)***
LOG(SIZE)	-0.0168		0.4142	
	(0.0005)		(1.065)	
PROPFIX	-0.9912		0.2335	
	(0.285)		(0.0357)	
P/E	-0.0065		-0.00624	
	(0.394)		(0.8352)	
MBV	-0.0513		-0.0255	
	(3.0916)*		(1.6144)	
GRDUMMY	-0.5515		0.1332	
	(0.6745)		(0.099)	
SepChair	-1.4627	-1.5927	-0.7817	-0.8671
	(4.9202)**	(8.416)***	(3.872)**	(5.974)**
NEX%	-0.1429		-0.0239	
	(0.5364)		(0.0254)	
NEXOutDir	-0.052		-0.0509	
	(0.4389)		(0.857)	
XoutDir	-0.3766	-0.3472	-0.0107	
	(3.9197)**	(5.136)**	(0.017)	
DirShHo	-0.0573	-0.0517	0.0139	
	(3.7107)*	(6.752)***	(0.9687)	
BigShHo	-0.00822		0.0245	
	(0.244)		(4.780)**	
Constant	7.2264	3.1119	0.7845	3.0077
	(3.8151)*	(16.9685)***	(0.154)	(10.967)***
N	104	104	156	156
Model chi-square	41.91***	25.73***	36.66***	22.965***
Likelihood Ratio Index	0.4175	0.2851	0.1929	0.1185
Correct classification %				
Target	82.69%	69.23%	67.95%	64.10%
Non-Target	82.69%	71.15%	73.08%	69.23%
Overall	82.69%	70.19%	70.51%	66.67%

Notes:

1 ***Significant at the 1% level, 2 tail test

**Significant at the 5% level, 2 tail test

*Significant at the 10% level, 2 tail test

2 Figures in the parentheses are Wald statistics.

3 The likelihood ratio index is defined as (1-log likelihood at convergence/log likelihood at zero). It is similar to the R^2 statistics of multiple regression and gives an indication of the Logit model's explanatory power.

All the above models are statistically significant at the 1% level. The model likelihood ratio index (0.4175) for TMD model is considerable higher compared with the other models. In TMD model, RSHF is statistically significant at the 1% level while TUT, SepChair, and XoutDir are at the 5% level, and GEAR, MBV, and DirShHo are at the 10% level. The overall classification rate is also very high. Unlike in MD model in chapter four, LIQD is statistically not significant. On the other hand, in comparison to MD model in chapter five, there is no difference. The significant financial variables indicate that target companies had lower returns on their shareholders' funds with lower growth and higher gearing. At the same time the corporate governance variables show that the roles of chairman and CEO are not separated and executive directors held less outside directorships in other companies while they held less shares in their companies compared with non-target companies.

The stepwise regression procedure on TMD left four variables. While XoutDir is statistically significant at the 5% level, RSHF, SepChair, and DirShHo are statistically significant at the 1% level. All the significant variables have their expected signs. The result shows that target companies are characterised as low return companies where management skill is graded relatively low in the labour market (XoutDir). It is also significant that target companies combine the roles of chairman and CEO, and target management interest is not aligned strongly with the shareholders compared to non-target companies (DirShHo).

The TTD is statistically significant at the 1% level with a likelihood ratio index of 19%. The variables, TUT, LIQD, SepChair and BigShHo are statistically significant at the five per cent level. The significant variables from the governance variables are SepChair and BigShHo. This again indicates that the target companies combine the roles of CEO and chairman and percentage of shareholding among the big shareholder is higher in target companies. It is interesting to note that this is exactly the same compared with TD models in financial ratio and corporate governance analysis. Even though it is not statistically significant the negative coefficient sign of MBV is in line with the expected sign indicating that the target company's shares are valued less in comparison with their book values, therefore lowering their MBV ratios. The other significant financial variables suggest that the target companies are characterised as having low growth and liquidity compared with the non-target companies in the sample.

A simple comparison of this model with TD in financial ratio analysis in chapter four shows that the overall classification percentages of the combined model is considerably higher as a result of better classification on the non-target companies. Also the likelihood ratio of TTD and TMD are higher which indicates that combined models provide a higher explanation of a firm's acquisition probability as a result of combined (as well as more characteristics included in the analysis) variables.

On the other hand, the stepwise regression of TTD provided three variables. These, as can be seen in Table 6.1, are the same significant variables, accept BigShHo, that were given by TTD, even though the statistical significance of 0.15 criterion for inclusion is used. The model itself is statistically significant at the one per cent level. TUT and LIQD, at the 1% level, and SepChair at the 5% level, are statistically significant variables provided by the stepwise elimination process. The result of the stepwise regression analysis suggests that during the time period outlined in this study the target companies were characterised by low growth and liquidity, and tended to combine the roles of chairman and chief executive directors on the boards. Nevertheless these significant differences can not be used in the identification of target companies prior to receiving bids.

The classification results of ANN for the above combined data samples are presented below.

Table 6.2. ANNs' classification table for TMD.15-6 nodes in the hidden layers

	Predicted		
Observed	Non-target	Target	%
Non-target	47	5	90.38%
Target	7	45	86.53%
		Overall	88.46%

An ANN with 15 nodes in the first and 6 nodes in the second layer provided the optimal network for the combined data and was used to test the network on the hold-out sample. The classification rates on the estimation sample are quite high with over 90% on the non-target companies and 86% on the target companies. The overall correct classification rate is slightly over 88%.

Table 6.3. ANNs' classification table for RMD. 6-4 nodes in the hidden layers

	Predicted		
Observed	Non-target	Target	%
Non-target	43	9	82.69%
Target	11	41	78.84%
		Overall	80.76%

For the RMD, the network with 6 nodes in the first and 4 nodes in the second layer gave the optimum network and was tested on the hold-out sample. The overall prediction accuracy rates of both networks are higher than LR classification rates on the estimation sample. The overall prediction rate has dropped to 80.76% with the RMD model.

Table 6.4. ANNs' classification table for TTD.15-5 nodes in the hidden layers

	Predicted		
Observed	Non-target	Target	%
Non-target	68	10	87.17%
Target	17	61	78.20%
		Overall	82.69%

An ANN with 15 nodes in the first and 5 nodes in the second layer provided the optimal network for the combined data and was used to test the network on the hold-out sample. The overall classification result for ANNs model is again higher than the one provided by logistic regression analysis.

Table 6.5. ANNs' classification table for RTD. 5-3 nodes in the hidden layers

	Predicted		
Observed	Non-target	Target	%
Non-target	60	18	76.92%
Target	25	53	67.94%
		Overall	72.43%

For the RTD, the network with 5 nodes in the first and 3 nodes in the second layer gave the optimum network and was tested on the hold-out sample. The overall prediction accuracy rates of both networks are higher than LR classification rates on the estimation sample. In the RTD the overall prediction rate has dropped from 82.69% in TTD to 72.43%. The network produced a higher classified rate for non-target companies compared with the target ones.

6.3.2 Hold-out Test Results

The prediction results of the above LR and ANNs models are presented below. Following the layout of previous chapters, the classification results of the estimated cut-off points as well as the 0.50 cut value will be displayed. Also, in the event that the overall classification of the models exceeds the proportional chance probability, the *t*-test for significance will be estimated. As it will be realised, when the estimated cut-off probability for reduced models is equal to 0.50, the predicted classifications will be equal to the ones with the total models' estimations at the 0.50 cut-off point.

6.3.2.1 Mixed Data Results

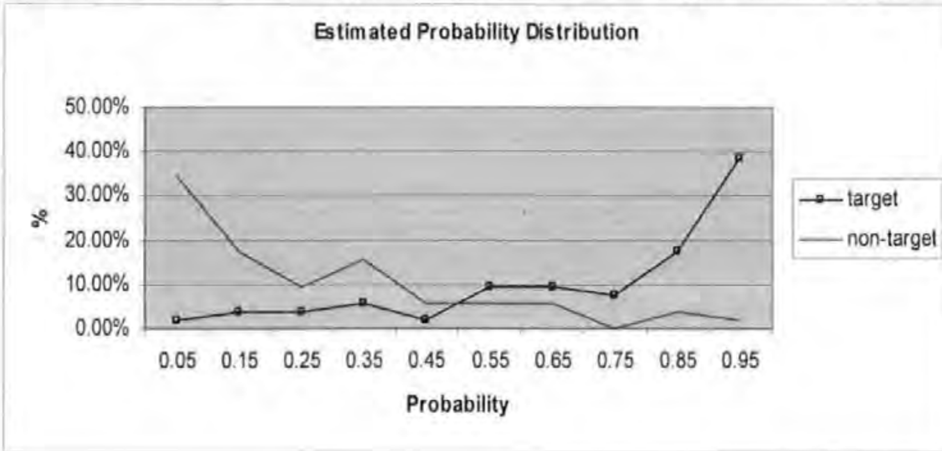
Table 6.6. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	84	61	57.93%
Target	20	31	60.78%
		Overall	58.67%

Table 6.7. Estimated distribution of targets and non-targets in the TMD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	1	1.92%	18	34.62%
0.100-0.199	0.15	2	3.85%	9	17.31%
0.200-0.299	0.25	2	3.85%	5	9.62%
0.300-0.399	0.35	3	5.77%	8	15.38%
0.400-0.499	0.45	1	1.92%	3	5.77%
0.500-0.599	0.55	5	9.62%	3	5.77%
0.600-0.699	0.65	5	9.62%	3	5.77%
0.700-0.799	0.75	4	7.69%	0	0.00%
0.800-0.899	0.85	9	17.31%	2	3.85%
0.900-0.999	0.95	20	38.46%	1	1.92%

Figure 6.1. Distribution of acquisition probabilities for targets and non-targets



The intersection between target and non-target probability occurs at 0.50 cut-off point. It can be seen that most of the target companies tent to have an acquisition probability greater than 0.5 while the non-target sample concentrates below the intersection point. From this distribution it can said that they show a very distinct and expected probability distribution.

At this cut-off point LR had an overall classification rate of 58.67% and below the proportional change criterion of 61.5%. This result is quite disappointing as the model has a very high explanation of acquisition likelihood. The model had a higher classification rate for targets compared with the non-target group even though the percentage difference is small. Percentage of targets classified correctly is 33.69% (31/92). This classification rate is not anyway different from the ones in chapter four and five.

Table 6.8.ANNs. (15-6 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	90	55	62.06%
Target	19	32	62.74%
		Overall	62.24%

ANNs prediction on the hold out sample is higher than LR classification as well as higher than could be obtained by the proportional change criteria of 61.5%. However the difference is not high enough to make it statistically significant. The t-statistic (0.2137) for this classification rate is not statistically significant. The correct classification rate of targets is 36.78% (32/87).

6.3.2.2 Reduced Mixed Data Results

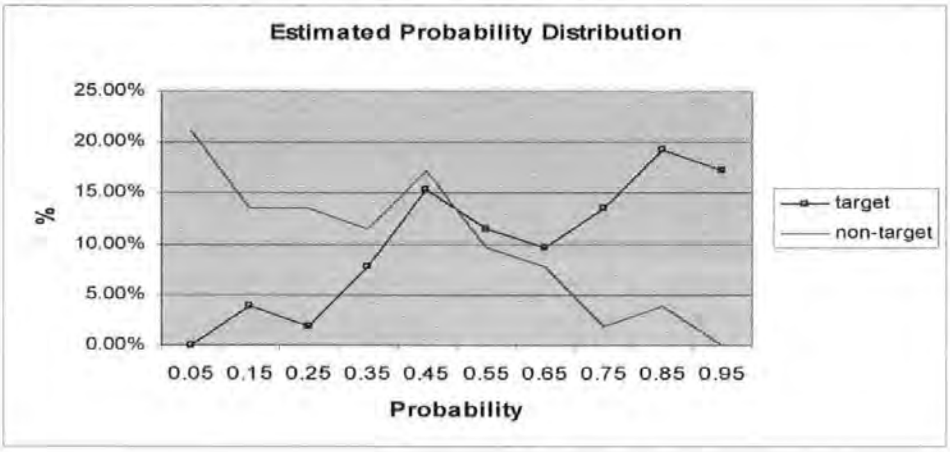
Table 6.9. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	84	61	57.93%
Target	20	31	60.78%
		Overall	58.67%

Table 6.10. Estimated distribution of targets and non-targets in the RMD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00%	11	21.15%
0.100-0.199	0.15	2	3.85%	7	13.46%
0.200-0.299	0.25	1	1.92%	7	13.46%
0.300-0.399	0.35	4	7.69%	6	11.54%
0.400-0.499	0.45	8	15.38%	9	17.31%
0.500-0.599	0.55	6	11.54%	5	9.62%
0.600-0.699	0.65	5	9.62%	4	7.69%
0.700-0.799	0.75	7	13.46%	1	1.92%
0.800-0.899	0.85	10	19.23%	2	3.85%
0.900-0.999	0.95	9	17.31%	0	0.00%

Figure 6.2. Distribution of acquisition probabilities for targets and non-targets



Again the two probability distributions intersect at 0.50 cut-off level. As a result the prediction rates are the same as the TMD.

Table 6.11.ANNs. (6-4 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	86	59	59.31%
Target	20	31	60.78%
		Overall	59.69%

The classification rate of the network is not very different from the LR classification rate. It has correctly predicted the same number of targets while it narrowly performed better on the non-target sample. The correct classification rate of targets is 34.44% and not different from LR classification rate.

6.3.2.1 Time Data Results

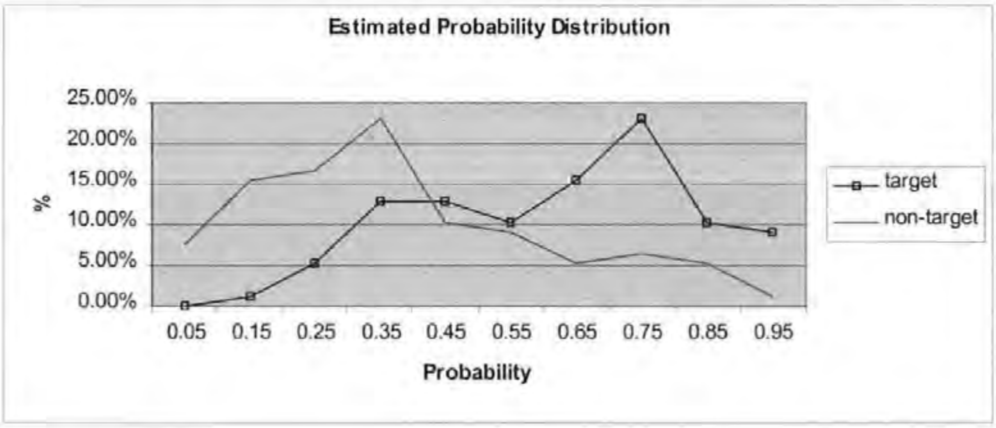
Table 6.4. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	89	30	74.79%
Target	11	14	56.00%
		Overall	71.53%

Table 6.13. Estimated distribution of targets and non-targets in the TTD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	5	6.41%	0	0.00%
0.100-0.199	0.15	6	7.69%	1	1.28%
0.200-0.299	0.25	12	15.38%	2	2.56%
0.300-0.399	0.35	9	11.54%	7	8.97%
0.400-0.499	0.45	18	23.08%	14	17.95%
0.500-0.599	0.55	14	17.95%	18	23.08%
0.600-0.699	0.65	8	10.26%	20	25.64%
0.700-0.799	0.75	4	5.13%	7	8.97%
0.800-0.899	0.85	2	2.56%	9	11.54%
0.900-0.999	0.95	0	0.00%	0	0.00%

Figure 6.3. Distribution of acquisition probabilities for targets and non-targets



As can be seen from the graph the probability distribution of the two groups intersects at 0.43. The classification of this probability point on the hold-out sample is presented in Table 6.14.

Table 6.14. LR. Cut-off value of 0.43

	Predicted		
Observed	Non-target	Target	%
Non-target	78	41	65.55%
Target	7	18	72.00%
		Overall	66.67%

The prediction rate of LR at the 0.43 cut-off point is lower than 0.50 cut value. The classification of target companies (72%) is considerably higher than at the cut value of 0.50 but at the same time it is notably smaller with the non-target ones. Furthermore, the percentage of actual target companies correctly classified within the group that is predicted as targets by the model is 30.51% (18/59) compared with 318% (14/44) of the 0.50 cut value. Even though the prediction rate for target companies is increased, the overall percentage classification is small. Furthermore, despite the fact that the overall prediction rate of LR at the cut-off point of 0.50 is above the proportional chance criterion of 71.3 per cent, the difference is not statistically significant.

Table 6.15.ANNs. (15-5 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	91	28	76.42%
Target	12	13	52.00%
		Overall	72.22%

The overall classification rate of the network is slightly higher than LR, as well as higher than would be expected by chance. The *t*-statistic for this rate is 0.246 and statistically not significant. The number of non-targets correctly classified by the network is slightly higher than the 0.50 cut value by LR. On the other hand, the network performed marginally lower with the target group, even though the difference is only one company. The percentage of actual targets within the predicted target group is 31.7% (13/41).

6.3.2.4 Reduced Time Data Results

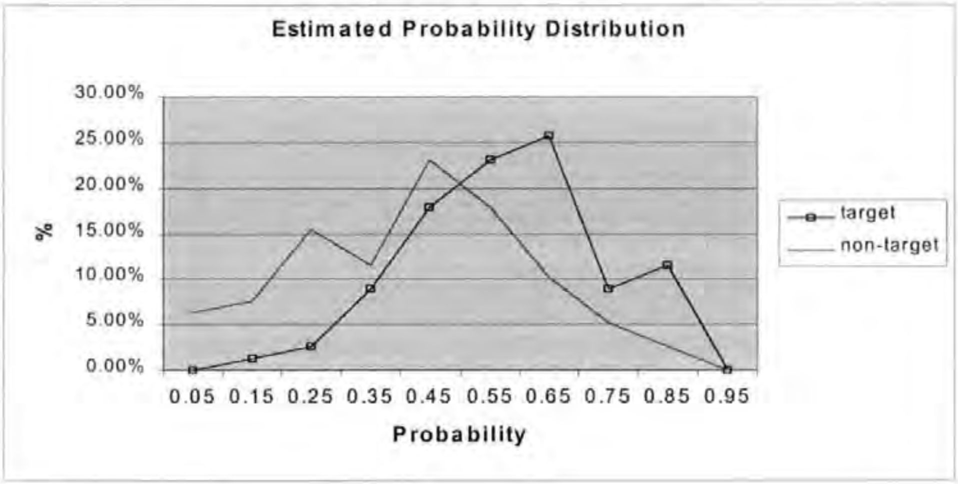
Table 6.16. LR. Cut-off value of 0.50

	Predicted		
Observed	Non-target	Target	%
Non-target	89	30	74.79%
Target	11	14	56.00%
		Overall	71.53%

Table 6.17. Estimated distribution of targets and non-targets in the RTD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00%	5	6.41%
0.100-0.199	0.15	1	1.28%	6	7.69%
0.200-0.299	0.25	2	2.56%	12	15.38%
0.300-0.399	0.35	7	8.97%	9	11.54%
0.400-0.499	0.45	14	17.95%	18	23.08%
0.500-0.599	0.55	18	23.08%	14	17.95%
0.600-0.699	0.65	20	25.64%	8	10.26%
0.700-0.799	0.75	7	8.97%	4	5.13%
0.800-0.899	0.85	9	11.54%	2	2.56%
0.900-0.999	0.95	0	0.00%	0	0.00%

Figure 6.4. Distribution of acquisition probabilities for targets and non-targets



The probability distribution of the two groups actually intersects at 0.49 and is very close to the 0.50 cut value. The classification of this probability point on the hold-out sample is presented in Table 6.11.

Table 6.18. LR. Cut-off value of 0.49

	Predicted		
Observed	Non-target	Target	%
Non-target	89	30	74.79%
Target	11	14	56.00%
		Overall	71.53%

The predictive result of the 0.49 cut value is identical to the 0.50 cut-off point. This is not surprising since these two values are very close to one another. LR classified 71.53% of the companies correctly. This is relatively a higher figure compared with the TTD (66.67%) above at the 0.43 cut-off point. On the other hand, the number of target companies correctly classified is down from 18 to 14 with 56%. Even though the overall classification is greater than would be expected from proportional chance probability, the difference is not big enough to be statistically significant.

Table 6.19.ANNs. (5-3 nodes)

	Predicted		
Observed	Non-target	Target	%
Non-target	94	25	78.99%
Target	13	12	48.00%
		Overall	73.61%

Although the network provided a slightly higher classification rate compared with LR, the result is higher than the proportional chance probability. However the *t*-statistic for this rate is 0.629 and not statistically significant. The network's classification of non-targets is higher than LR, even though it slightly underperformed with the target companies. The percentage of actual targets within the group that is predicted as targets is 32.43% (12/37) compared with 31.81% (14/44) in LR.

6.4 Discussions and Summary

In this chapter, in order to further the prediction abilities of the two techniques, financial and corporate governance variables of the companies in the sample are combined. This method was only applied to mixed and time data and the underlying reasons for this already mentioned above. Also, a stepwise regression analysis with the stepwise elimination method was applied and tested on the hold-out sample. The estimated LR models are statistically significant at the one per cent level.

The stepwise regression result of TMD suggests that four variables (RSHF, SepChair, XoutDir, and DirShHO) are contributing significantly to the separation of target companies from the matched sample of non-targets. It is interesting to note that among these four only one (RSHF) is a financial variable while the rest of them came from non-financial characteristics. On the financial side the estimation sign of return on shareholders' fund indicates that target companies had provided lower returns to their shareholders. Furthermore on the governance side target companies combined the positions of chairman and CEO that put too much power in the hands of one person, their executive directors had lower outside directorships indicating that their managerial skills were not so highly rated by the managerial labour market compared with non-target companies, and their directors held less shareholding in their companies leading to misalignment of their interest with the shareholders.

On the other hand the stepwise regression analysis of TTD produced three significant variables (TUT, LIQD, and SepChair) and only one of them (SepChair) is still statistically significant compared with RMD. The results of RTD suggest that target companies are characterised as low growth and low liquidity. However the growth-resource dummy is not significant and does not suggest any growth-resource imbalance. But the significance of TUT and LIQD indicates that target companies had lower growth with liquidity problems. Furthermore the separation of the duties between chairman and CEO is still significant and the negative sign demonstrates that target companies combined these roles.

Although the models provided a statistically significant explanation of a firm's acquisition probability, the classification results of the models on the hold-out sample did not provide significantly higher prediction rates. Neither ANNs nor LR supplied significantly higher prediction rates above the proportional chance probability of 71.3% on the time data. On the other hand, ANNs had a classification rate above the proportional chance probability of 61.5% on the TMD but this was not statistically significant either. Once again ANNs failed to outperform LR results. Although the network's overall classification rates were higher than LR, the differences were not significant.

As in the previous chapters, the results obtained from the models above show that it is not possible to predict target companies prior to receiving bids by using the above (financial and non-financial) information. However, the results are limited to the variables used and the techniques applied.

Chapter 7 - Conclusions

7.1 Introduction

As summarised earlier, the previous research suggests that takeovers are value enhancing and target companies' shareholders benefit from such events as premiums are paid above the market value. This has lead to a quest to identify target companies before they receive a bid as it creates an opportunity to obtain abnormal returns, if they are predicted. Although a substantial amount of research is dedicated to the predictions of takeover targets and claims of successful modelling, it has not actually been possible to predict takeover targets in advance of the stock market. These studies are already identified and quoted in the earlier chapters.

In the previous takeover prediction research, not only different characteristics of companies are included in the modelling process but also different statistical and multivariate techniques are employed as modelling tools. Although these techniques cover a wide range, the most popular of these are logistic regression, multiple discriminant analysis and probit analysis. Recently, application of artificial neural networks in other areas of finance has grown exponentially. The main applications of them in prediction modelling are in time series and corporate failure predictions. In terms of its methodology corporate failure predictions are similar to takeover predictions and reported successful application of ANNs in this area prompted the employment of this technique in takeover predictions as it has never been applied, at least to corporate governance characteristics of target companies.

Therefore in this research, ANNs were applied to takeover predictions along with Logistic Regression. Two types of company data were used during the analysis. First, the two techniques were put to the test by employing financial ratios of the companies and second their predictive abilities were evaluated with corporate governance variables. In these analyses data relating to one and two years prior to takeovers, as well as the percentage change of this data (variation data) from two years to one year were extracted for two estimation samples (mixed and time data) and tested on the respective hold-out samples.

In chapter four, the modelling of takeover predictions was conducted by using financial ratios as financial variables. The ratios were selected according to the relevant hypotheses that were specified in the literature as significant reasons for acquisitions. The selection of each ratio within the outlined hypothesis was similar to that used by Palepu (1986) with one additional ratio used by Ambrose and Megginson (1992). The reason for selecting similar ratios to those used by Palepu (1986) was to use them as a yardstick for measurement and comparability since the same financial ratios were used by Sen *et al.*, (1995).

In chapter five, the corporate governance variables of the sample companies used to discriminate between the target and non-target companies was tested on a hold-out sample for prediction purposes. The six variables were selected as measures of what is perceived to be an efficient internal monitoring mechanism in the previous literature and by the Cadbury Committee. As explained previously, the reasons for using corporate governance variables were twofold; first the growing literature on the importance of an optimal corporate governance structure on firm performance (Shivdasani, 1993; Morck *et al.*, 1988), and of the board characteristics of target companies (Gammie and Gammie, 1996; Weir 1997).

Finally, in chapter six, the two sets of variables were combined to test whether this would lead to an increase in the predictive abilities of the previous models. Apart from using one of the models in the prediction, a stepwise regression analysis was also carried out to eliminate insignificant variables from the model.

Although this research relates, in terms of its methodology and application of similar financial variables, to Palepu's (1986) influential study in takeover predictions, it differs from it by the application of ANNs. This study also differs from the analysis of Sen*et al.*, (1995), by its application of non-financial characteristics of the companies along with the combined (financial and non-financial) data.

7.2 Empirical Results and General Remarks

The reported empirical results in chapter four, five and six show that neither ANNs nor LR produced statistically significant prediction results in the identification of potential targets. Also contrary to the reported success of ANNs in comparison with LR in corporate failure prediction studies (Alici, 1996), ANNs did not provide better prediction results with either type of data compared to LR. In general, the only prediction results that were above the proportional change probabilities were those with the financial ratios of mixed data. ANNs and LR achieved 64.28 and 62.76 per cent overall classification rates which were above what could be obtained by chance. The *t*-statistics for these prediction rates were 0.364 for LR and 0.812 for ANNs. However, these statistics are not statistically significant. Also, as mentioned, this type of sample was only formulated to act as a benchmark to the time data modelling. However, even though ANNs failed the expectations of a significant target prediction, it is shown that they are extremely capable of providing mathematical fit to complex non-linear data.

Despite the fact that financial ratio modelling supplied a higher explanation of takeover probabilities of firms compared with governance data, the prediction results of these models were not significantly different. TD(-1) modelling in financial ratio analysis was significant at the one per cent level with a likelihood ratio of 0.136 compared with a lower model chi-square and likelihood ratio of 0.070 with the TD(-1) modelling in governance data. However, the overall prediction rates of both data were not higher than the proportional chance criterion and not statistically significant.

The effort to improve the predictive abilities of the models by combining the two sets of data did not provide better prediction results either. Although the estimated model (TTD) was statistically significant at the one per cent level with 0.1929 likelihood ratio, the overall prediction rates were above the chance probability but not statistically significant. In this type of data, on the other hand, ANNs' classification was slightly higher compared with LR. Also a stepwise regression analysis was applied to the combined time data and found that three variables -TUT, LIQD, and SepChair- outlined the significant differences between the target and non-target companies one year prior to receiving a takeover bid. The result of the stepwise regression analysis suggests that, during the time period outlined in this study, the target companies were characterised by low growth and liquidity, and tended to combine the roles of chairman and chief executive directors on the boards. Nevertheless, these significant differences can not be used in the identification of target companies prior to receiving bids.

The predictive results of time data modelling suggest the time impact of takeover likelihood characteristics of targets. In short, a model formed with past data would not yield significant insight to the following periods to be used in the prediction of target companies by employing publicly available information. This characteristic of takeovers is actually one of the causes of failure in successful prediction of targets. Whether this prediction is done by LR or ANNs. As stated, ANNs learns from experience, from the data in this case, and projects this learning onto the test data. The networks' learning, however, is limited to the experience gained from derivation data and if the same pattern is not to be found in the test data it is obvious that the prediction will be incorrect.

The results are consistent with the semi-strong form of the efficient market hypothesis which states that current security prices instantaneously and fully reflect all publicly available information about securities markets (Blake 1990). Therefore, by simply using publicly available financial information that is already reflected in the share price, it is not possible to beat the market. This was also confirmed in recent research by Barnes (1998), but it was reported that the anticipatory price changes might indicate that the market may not be efficient in the strong form.

The high success rates that have been achieved in corporate failure studies stems from the fact that acquisitions and corporate failures differ substantially. It is easier to predict eventual failure if the firm has been consistently making losses, not generating enough liquid funds, losing its market share to its competitors.

7.3 Limitations of the Study

The first limitation of this study is the variables that were used in the modelling process. In financial ratio modelling, for example, the variables are selected from the postulated hypothesis of takeovers and in order to relate them to Palepu's (1986) study, similar ratios are employed in this study. The number of ratios employed can be extended with the inclusion of other ratios. As mentioned above, one of the most important characteristics of takeovers is that they vary in time so that it should be expected that the different ratios, as a result of changing motives for bidders and factors of the market, become important at different times. Also, the representative ratio used of each hypothesis may not be the best one. A group of ratios under each hypothesis can be employed in the modelling process. Barnes (1998), for example employed 18 ratios altogether in his analysis and under the inefficient management hypothesis there were seven ratios. There is of course the danger of creating multicollinearity in the estimation data which would lead to bias in the estimated model. Second, the variables in this study are based on historical cost data and employment of current cost data may improve the discrimination ability of the models. Walter (1994), reported that "employing current cost data was useful for identifying future takeover targets and earning above-average stock returns"(p.349). Therefore, the prediction ability of the models might be improved with current cost data or with other ratios.

For the governance data, the employed variables can be extended for the better measurement of companies' internal monitoring process. These could include, for example, the share options held by the executive directors or the identification of the affiliation of large shareholder and non-executive directors to the executive directors. Shivdasani (1993), found that ownership by management with unaffiliated large shareholders decreases the probability of takeover, while ownership by unaffiliated large shareholders increases this probability.

There are various other factors that come into effect to have an impact on the acquisition probability of a company. Takeover defence measures that are employed by companies significantly affect their acquisition probability. Sudarsanam (1991) reports different takeover defence strategies employed by UK firms. It can be seen in his study that pre-bid defensive strategies are solely employed to deter potential bidders. These strategies increase the cost of acquisition and reduce the expected benefits of the acquisition to bidders, hence reducing the takeover likelihood of companies. Along with these strategies, for example, employee share ownership plans (ESOP) can be used as an entrenchment tool by the incumbent management even though they cannot be used as a poison pill under UK company law (Sudarsanam, 1995, p.200). ESOPs can also have an indirect effect on the acquisition probability of the firms. Park and Song (1995) found significant improvement in year-end performance of the companies that employed ESOPs.

Also identifying the nature of the takeover bid as friendly or hostile could improve the prediction ability of the models. Takeover characteristics of firms may differ according to the nature of the bids (Powell, 1997). It is likely that hostile takeovers are more disciplinary than friendly ones (Shivdasani 1993). However Franks and Mayer (1996) reported that hostile bids are not directed to poorly performing firms in the UK. This results in different motives of bidders, as a result affecting the characteristics of targets. Weir's (1997) analysis of corporate governance data shows some of the differences between the hostile and non-hostile characteristics of target companies. The study reports that the percentage shareholding of executive directors is statistically significant at the 10% level in the hostile group, while this is not significant in the non-hostile group. However due to the nature of the acquisition market in the UK, the number of hostile takeovers are smaller than in the US market. Sudarsanam (1991), Weir (1997), and Powell (1997) reported that the proportion of hostile takeovers during the 1980s and 1990s is around 22 to 24 per cent in the UK.

Furthermore the construction of the hold-out sample in this study is biased. First, the sample does not reflect the population characteristics where the companies are not divided into two groups of target and non-target companies. During the forming of the sample, special attention has been given to ensure that the non-target companies had not received any takeover bids two years before and after the time period analysed and had their complete data. This process obviously excluded some companies and reduced the number of non-target companies included in the sample and, as a result, led to a different data composition than the population. Some companies were not selected as non-target companies in the sample because they disappeared from the population for reasons other than being taken over. This process of selecting target companies enhances the differences between target and non-target firms in the estimation sample and would result in better prediction in the test sample than could be expected in the population.

On a more technical level, in the application of ANNs, the analysis of internal connections in the network has not been attempted. This process could lead to the identification of important features in the network connections as well as the determination of the best-input variables (Vaughn 1996). Such improvements would lead to the identification of best ratios.

Some other impediments of takeover prediction modelling in general can be summarised as:

In theory the matching of the companies in this type of empirical research implies that there are two groups of subjects which are, as groups, similar in all the aspects of the business that they are operating in but different in certain characteristics that lead to a takeover bid. However, the restrictions or impediments attached to the data can quite easily prevent any prediction study from forming a clear cut classification or distinction of the groups. In most of the prediction studies, as in this study, matching is carried out on the basis of very broad industry classification. In any industry a company may receive a bid and the underlying reasons for it may not come from a target's financial or non-financial characteristics but may simply be as a result of the bidder's characteristics or intentions. Powell (1997) reported that models of takeovers based on the characteristics of targets only might not contain enough information. Tzoannos and Samuels (1972) also reported that "*...linear probability function estimated for companies bidding for others has a better explanatory power than the one estimated for companies taken over*" p.15. Sudarsanam (1995) indicate that acquisitions should be examined in the context of a company's broader business and corporate strategy. Each acquisition type is a result of these business and corporate strategies and the acquisition type dictates the target profile and assessment of target companies. In addition to that, for example, in an environment of an Endgame theory, as defined by Martin (1997), managers start acquiring other companies in order to prevent themselves from being acquired. In this sort of environment, it is obvious that they are not pursuing shareholders' wealth maximisation but wish to acquire other companies before being acquired. Therefore, inability of any model to incorporate aims or intentions of the bidding companies is a major limitation of any takeover prediction study.

Even an attempt to include every possible characteristic of target companies along with target and bidding companies to find out reasons why takeover bids make economic sense or to discover the synergy, the models may not discriminate between two groups of companies because there may not be any financial or economic logic in the event and, as a result, any characteristics to differentiate them. Roll (1994) already argued that corporate takeovers are one area that is separate from the rational behaviour of aggregate market behaviour but characterised by irrational behaviour of individuals as a result of hubris and valuation error. Another factor to consider is that de-mergers and divestment activities happen after the merger or takeover. One also should not ignore the fact that information efficiency in the financial markets does not necessarily imply the same efficiency in the other areas related to the targets such as marketing strategies and production technologies (Aggarwal and Navratil, 1991).

Relying on an efficient market and efficient management in such circumstances as endgame theory does not make any sense. For example, the market for corporate control is considered to be unaffected by the social relations and described as an asocial conceptualisation. However, managerial actions are influenced by the current social structures and are not determined entirely by economic interests. Inter-organisational structures affect the perception of interest and the manner in which interests are pursued (Davis, 1991).

As mentioned before, the reasons for takeovers cover a wide spectrum of motives and stem from the bidders' point of view and these are the ones that have been identified in the literature as efficient corporate objectives. Managers of the bidding firms look at, let us assume, financial ratios in this simplistic case, specific financial ratios that satisfy the company objectives. Classifying companies as potential targets makes sense for bidders since their strategy for acquisition has been established and finding the suitable target is a matter of simple ranking or classification. Asquith *et al.* (1983) drew attention to the acquisition programmes. If one also tries to take into account other self-motivated objectives by bidding management, the reflection of these objectives in the targets' financial ratios that are being used to discriminate will dissipate among all the targets' ratios and using these ratios to discriminate will be fruitless. Unlike corporate failure predictions, where whatever the reason for going bankrupt there is and will be an identifiable pattern in the ratios of the bankrupt firms, there will not be a clearly formed pattern in the case of takeover targets (Singh, 1975).

Mergers are a method of investment and require a proper investment appraisal assessment. A bidder is involved in such a process if the expected incremental cash flow from the merged firm, after discounting for the cost of capital which incorporates consideration of market perceptions of the risk of the merged firm, is positive. To put it simply, if there is an opportunity to achieve a positive net present value (NPV) (McLaney, 1997).

NPV assessment of an investment opportunity involves the future costs and revenues of the combined firm where the bidder's valuation of target plays a crucial role. "*The value of the target from the bidder's point of view is the sum of the per-bid stand-alone value of the target and the incremental value the bidder expects to add to the target's assets*" (Sudarsanam, 1995, p.138). If the bidder's valuation of a target is bigger than the market's valuation of the target the bidder makes a bid. In contrast to the market valuation of a company, which is the stock market value of a firm's equity capital, the bidder's valuation of a target is subjective, which includes any estimated gains through synergy and weak management, and cannot be observed (Roll, 1994). Basically, the bidder firm makes an assessment and a bid for a target because it estimates that the earning power of the target's assets will be enhanced under the new management. Since it is not possible to observe the bidder's valuation of the target and see why they estimate that their valuation is higher than the market's, it is not easy to predict takeover targets by simply looking at targets characteristics without a proper knowledge and background of the bidder's intentions.

The implicit assumption about takeover prediction originates from bankruptcy prediction and assumes a distinctive recognisable pattern between targets and non-targets without considering and including the bidders in the equation. This would only be valid if a single reason which could have explained mergers as a means of restructuring distressed firms (Clark and Ofek, 1994). As mentioned above, information efficiency is restricted to the financial information and financial information does not incorporate production technologies of the targets. If a target's production technology or marketing niche is what a bidder is after it may not pay too much attention to the target's ratios and pay some excess price for it, regardless of the hubris that may possess the bidding management.

Another difficulty arises from the fact that takeovers are a dynamic process where the impact on takeover likelihood characteristics of targets change over time (Powell, 1997). This problem is recognised as the time-series problem by Altman *et al.*, (1981). The time series problem occurs in takeover prediction studies because takeovers occur too infrequently within a year to generate enough target sampling. This requires pooling of different time periods that incorporate structural changes. This then effects not only the variables but also the priori probabilities. As a result, the identified characteristics of targets sampled over a long period of time will not result in a consistent pattern and even those that are possible to detect will diffuse. Choosing a shorter time period to eliminate this disadvantage will result in a better recognition of these characteristics, this is again assuming that in this short period time, those different objectives of the bidders also converge due to, for example, the market induced conditions. However, it will be very difficult to use the model for future takeover predictions. Also one should not exclude the importance of the legislative environment on mergers and acquisitions (Bartley, 1990; Sudasarnam 1995).

7.4 Further Research

The approach followed in this study can be extended in various ways to further research. In the case of ANNs, one should experiment with different activation functions. Although the logistic sigmoid function is used in this study, other similar functions that were mentioned in chapter 3 could be used. Sen *et al.*, (1995) for example a applied hyperbolic tangent activation function. It could be argued that the similar classification result achieved by ANNs and LR is as a result of employing a logistic sigmoid function in this study. After all, the maximum likelihood method (MLM) in LR chooses that estimators of the set of unknown parameters, which maximises the likelihood of observing the data that has been collected (Kleinbaum 1994). MLM is an iterative process. Therefore, employing a standard back propagation algorithm (SBP) with a logistic sigmoid function in ANNs is a broadly similar process to LR and bound to produce similar results.

The optimum brain damage (OBD) (Cun *et al.*, 1990) or Skeletonization process (Alici, 1996) can be used to extract important variables from the network through a continuous stepwise elimination. This approach will not only provide an optimum network structure but the training time through fewer training examples and an improved speed of learning (Cun *et al.*, 1990; Mozer and Smolensky, 1989). These variables may however not be very different from the ones produced by the stepwise regression process applied in this study. Hybrid and complex networks can also be applied (Maren, 1990).

ANNs are a sub-discipline of Artificial Intelligence (AI) and other related AIs can be applied to takeover prediction studies. Genetic Algorithms can be used to formulate the optimum network structure. Neuro-fuzzy nets for example extract the relationship from data set. The relationship from a given data is extracted by neural networks and the rules are set by fuzzy reasoning. Also the actual model created by the net can be analysed. This would enable the researcher to examine the model in detail. This allows for a greater understanding of how the model operates.

Although the models formed in chapters four and five show that the two years prior to takeover data did not provide a significant insight into the differences between target and non-target companies, the models that were formed with variation data seemed to be more robust. Furthermore, the difficulties of making inferences about the dynamics of chance from cross-sectional analysis have already been reported by Hermalin and Weisbach (1991) in the effect of board composition on firm performance and Hsiao (1986) in economic research. Therefore, panel data analysis could be used. In this way, more complicated models can be constructed (Hsiao, 1986).

7.5 Summary

The arguments presented above suggest that the takeovers, especially the reasons for companies to acquire specific targets manifest extremely complex behaviour that changes over time and cannot be modelled by applying ANNs and LR, at least with the variables used in this study, to predict future target companies for the purpose of gaining an excess return in the market.

As quoted by Young, (1999), "in a recent Centre for the Study of Financial Innovation round table meeting on competitive/competitor intelligence in the UK, some of the speakers said that artificial intelligence is just that, artificial. And there is no substitute for the human brain. One participant said that looking for anomalies is an art form: intuition is something that cannot be recreated by a computer. He also said that neural networks use past data (more emphasis should be on current data), and points to the disastrous effects that came about in the near-collapse of hedge fund Long-Term Capital Management as a prime example of too much reliance on computers"(p65).

Appendices

APPENDIX A

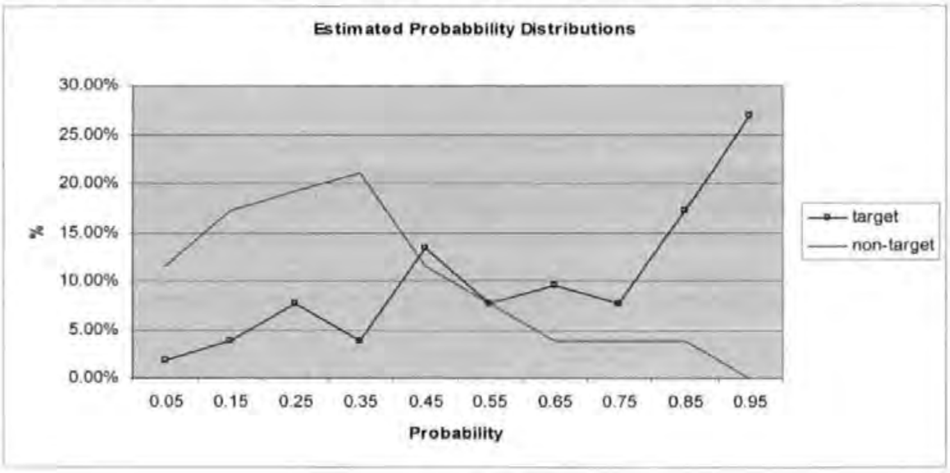
DATA FOR THE ESTIMATION OF CUT-OFF PROBABILITIES - FINANCIAL DATA

A.1 Probability distributions of targets and non-targets in the Mixed Data (MD).

Table A.1.1. Estimated distribution of targets and non-targets in the MD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	1	1.92	6	11.54
0.100-0.199	0.15	2	3.85	9	17.31
0.200-0.299	0.25	4	7.69	10	19.23
0.300-0.399	0.35	2	3.85	11	21.15
0.400-0.499	0.45	7	13.46	6	11.54
0.500-0.599	0.55	4	7.69	4	7.69
0.600-0.699	0.65	5	9.62	2	3.85
0.700-0.799	0.75	4	7.69	2	3.85
0.800-0.899	0.85	9	17.31	2	3.85
0.900-0.999	0.95	14	26.92	0	0.00

Figure A.1.1. Distribution of acquisition probabilities for targets and non-targets

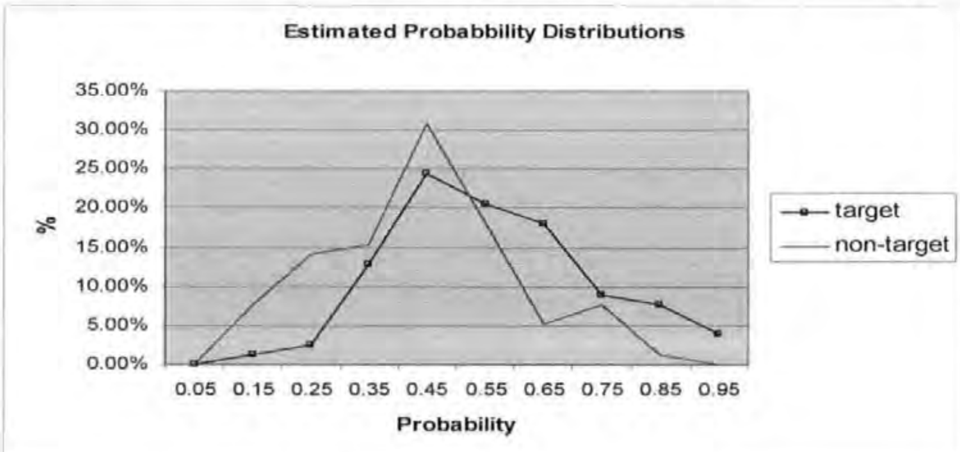


A.2 Probability distributions of targets and non-targets in the Industry Relative Data (IRD).

Table A.2.1. Estimated distribution of targets and non-targets in the IRD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	0	0.00
0.100-0.199	0.15	1	1.28	6	7.69
0.200-0.299	0.25	2	2.56	11	14.10
0.300-0.399	0.35	10	12.82	12	15.38
0.400-0.499	0.45	19	24.36	24	30.77
0.500-0.599	0.55	16	20.51	14	17.95
0.600-0.699	0.65	14	17.95	4	5.13
0.700-0.799	0.75	7	8.97	6	7.69
0.800-0.899	0.85	6	7.69	1	1.28
0.900-0.999	0.95	3	3.85	0	0.00

Figure A.2.1. Distribution of acquisition probabilities for targets and non-targets

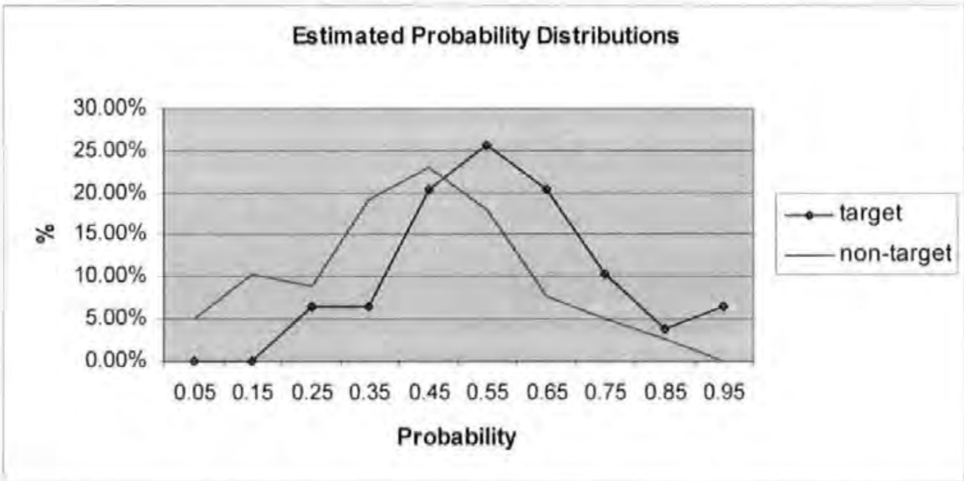


A.3 Probability distributions of targets and non-targets in the Variation Data (VD).

Table A.3.1. Estimated distribution of targets and non-targets in the VD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	4	5.13
0.100-0.199	0.15	0	0.00	8	10.26
0.200-0.299	0.25	5	6.41	7	8.97
0.300-0.399	0.35	5	6.41	15	19.23
0.400-0.499	0.45	16	20.51	18	23.08
0.500-0.599	0.55	20	25.64	14	17.95
0.600-0.699	0.65	16	20.51	6	7.69
0.700-0.799	0.75	8	10.26	4	5.13
0.800-0.899	0.85	3	3.85	2	2.56
0.900-0.999	0.95	5	6.41	0	0.00

Figure A.3.1. Distribution of acquisition probabilities for targets and non-targets

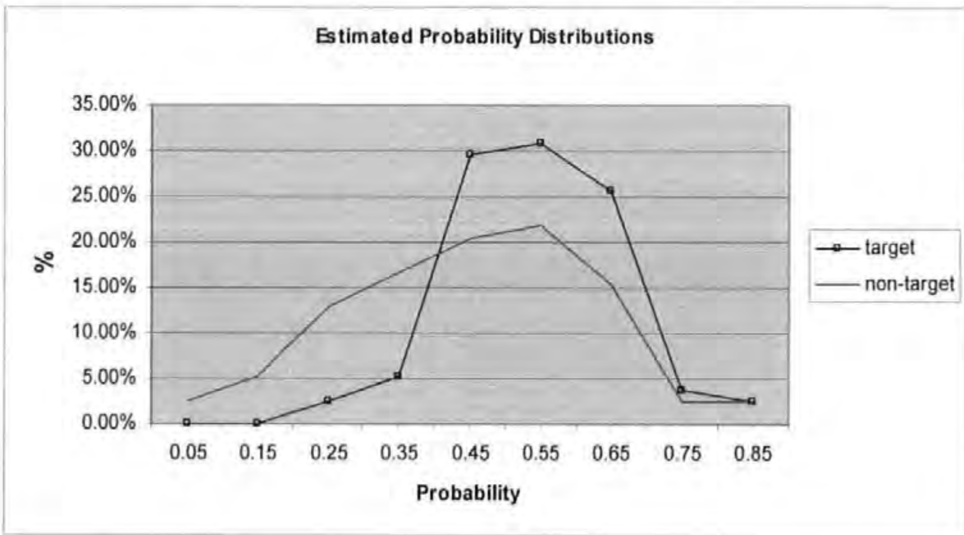


A.4 Probability distributions of targets and non-targets in the TD (-2).

Table A.4.1. Estimated distribution of targets and non-targets in the TD(-2)

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	2	2.56
0.100-0.199	0.15	0	0.00	4	5.13
0.200-0.299	0.25	2	2.56	10	12.82
0.300-0.399	0.35	4	5.13	13	16.67
0.400-0.499	0.45	23	29.49	16	20.51
0.500-0.599	0.55	24	30.77	17	21.79
0.600-0.699	0.65	20	25.64	12	15.38
0.700-0.799	0.75	3	3.85	2	2.56
0.800-0.899	0.85	2	2.56	2	2.56
0.900-0.999	0.95	0	0.00	2	2.56

Figure A.4.1. Distribution of acquisition probabilities for targets and non-targets



APPENDIX B

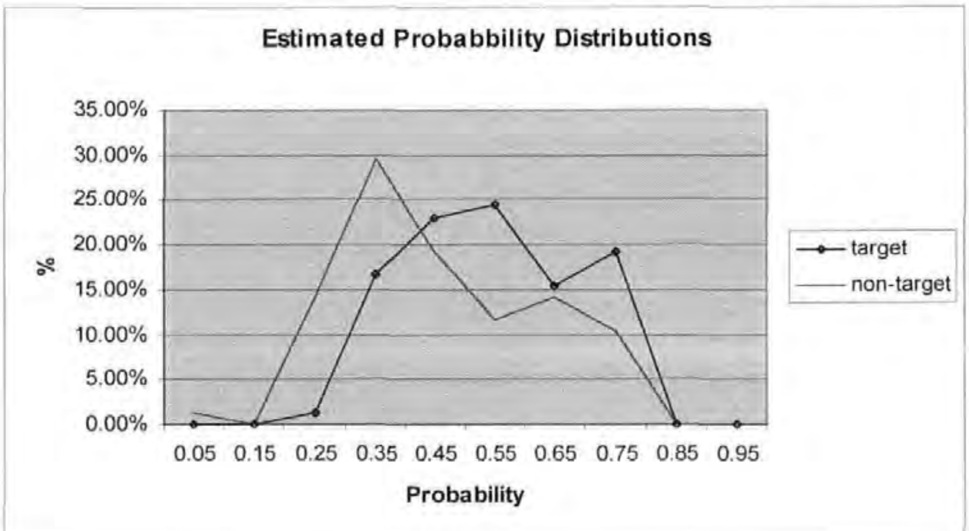
DATA FOR THE ESTIMATION OF CUT-OFF PROBABILITIES
NON-FINANCIAL DATA

B.1 Probability distributions of targets and non-targets in the Time Data (-1).

Table B.1.1. Estimated distribution of targets and non-targets in the TD (-1).

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	1	1.28
0.100-0.199	0.15	0	0.00	0	0.00
0.200-0.299	0.25	1	1.28	11	14.10
0.300-0.399	0.35	13	16.67	23	29.49
0.400-0.499	0.45	18	23.08	15	19.23
0.500-0.599	0.55	19	24.36	9	11.54
0.600-0.699	0.65	12	15.38	11	14.10
0.700-0.799	0.75	15	19.23	8	10.26
0.800-0.899	0.85	0	0.00	0	0.00
0.900-0.999	0.95	0	0.00	0	0.00

Figure B.1.1. Distribution of acquisition probabilities for targets and non-targets.

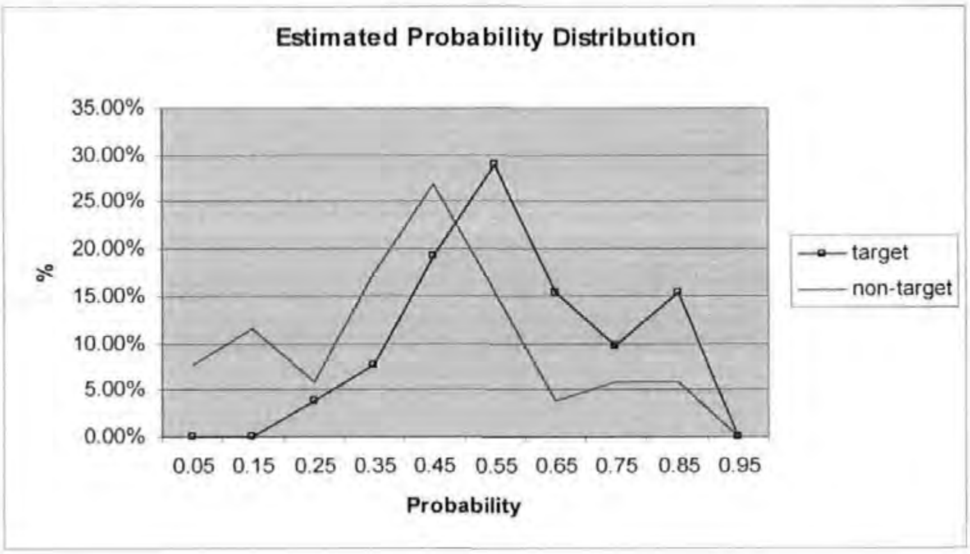


B.2 Probability distributions of targets and non-targets in the Mixed Data (MD).

Table B.2.1. Estimated distribution of targets and non-targets in the MD.

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	4	7.69
0.100-0.199	0.15	0	0.00	6	11.54
0.200-0.299	0.25	2	3.85	3	5.77
0.300-0.399	0.35	4	7.69	9	17.31
0.400-0.499	0.45	10	19.23	14	26.92
0.500-0.599	0.55	15	28.85	8	15.38
0.600-0.699	0.65	8	15.38	2	3.85
0.700-0.799	0.75	5	9.62	3	5.77
0.800-0.899	0.85	8	15.38	3	5.77
0.900-0.999	0.95	0	0.00	0	0.00

Figure B.2.1. Distribution of acquisition probabilities for targets and non-targets.

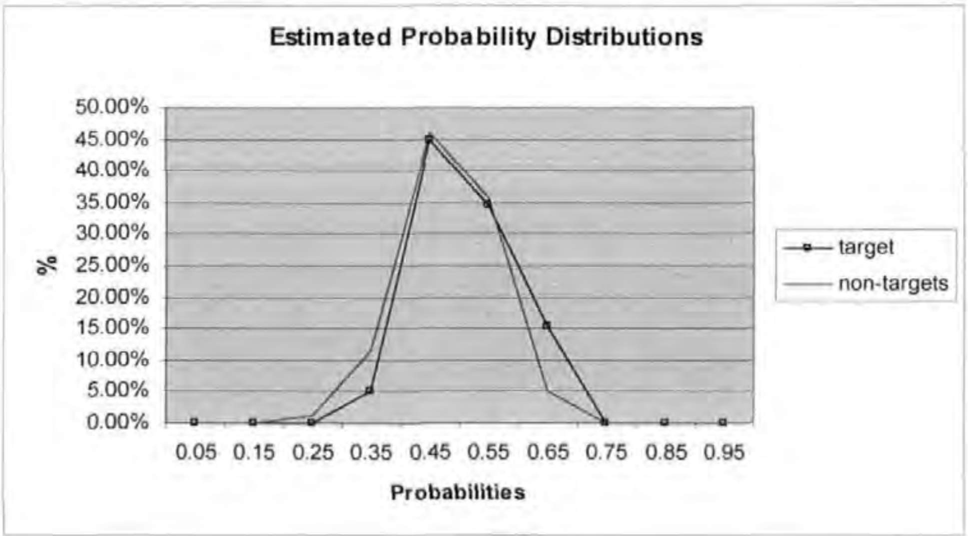


B.3 Probability distributions of targets and non-targets in the Time Data (-2).

Table B.3.1. Estimated distribution of targets and non-targets in the TD (-2).

		Target		Non-Target	
Range	Mid Value	No	%	No	%
0.000-0.099	0.05	0	0.00	0	0.00
0.100-0.199	0.15	0	0.00	0	0.00
0.200-0.299	0.25	0	0.00	1	1.28
0.300-0.399	0.35	4	5.13	9	11.54
0.400-0.499	0.45	35	44.87	36	46.15
0.500-0.599	0.55	27	34.62	28	35.90
0.600-0.699	0.65	12	15.38	4	5.13
0.700-0.799	0.75	0	0.00	0	0.00
0.800-0.899	0.85	0	0.00	0	0.00
0.900-0.999	0.95	0	0.00	0	0.00

Figure B.3.1. Distribution of acquisition probabilities for targets and non-targets.



APPENDIX C

CORRELATION MATRIXES - FINANCIAL DATA

C.1 Correlation Matrix - Mixed Data (MD).

Obs	_TYPE_	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY
1	MEAN		0.5	15.075	7.12	67.605	4.476	0.996	0.381	12.557	12.023	0.327
2	STD		0.502	41.6	18.327	68.473	0.636	0.232	0.184	38.393	10.477	0.471
3	N		104	104	104	104	104	104	104	104	104	104
4	CORR	DV	1	-0.369	-0.325	0.198	0	-0.215	0.035	-0.101	-0.246	-0.041
5	CORR	RSHF	-0.369	1	0.369	-0.062	0.313	0.056	0.016	0.012	0.207	0.011
6	CORR	TUT	-0.325	0.369	1	0.015	0.159	-0.152	0.021	-0.102	-0.01	-0.044
7	CORR	GEAR	0.198	-0.062	0.015	1	0.191	-0.326	0.085	-0.195	-0.014	-0.052
8	CORR	LOG(SIZE)	0	0.313	0.159	0.191	1	0.028	0.429	0.152	0.145	-0.235
9	CORR	LIQD	-0.215	0.056	-0.152	-0.326	0.028	1	-0.081	-0.004	0.024	0.236
10	CORR	PROPFIX	0.035	0.016	0.021	0.085	0.429	-0.081	1	0.1	-0.08	-0.053
11	CORR	P/E	-0.101	0.012	-0.102	-0.195	0.152	-0.004	0.1	1	0.303	-0.073
12	CORR	MBV	-0.246	0.207	-0.01	-0.014	0.145	0.024	-0.08	0.303	1	0.06
13	CORR	GRDUMMY	-0.041	0.011	-0.044	-0.052	-0.235	0.236	-0.053	-0.073	0.06	1

C.2 Correlation Matrix - Time Data (-1).

Correlation Matrix - TD (-1)												
Obs	_TYPE_	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P_E	MBV	GRDUMMY
1	MEAN		0.5	14.295	5.585	67.947	4.418	1.005	0.371	9.567	11.299	0.314
2	STD		0.502	36.563	22.421	57.37	0.621	0.229	0.165	33.668	9.997	0.466
3	N		156	156	156	156	156	156	156	156	156	156
4	CORR	DV	1	-0.265	-0.267	0.159	-0.011	-0.173	0.074	-0.098	-0.148	0.014
5	CORR	RSHF	-0.265	1	0.502	-0.043	0.236	0.121	-0.125	0.15	0.27	0.042
6	CORR	TUT	-0.267	0.502	1	-0.057	0.038	-0.049	-0.019	-0.052	0.102	-0.139
7	CORR	GEAR	0.159	-0.043	-0.057	1	0.139	-0.409	0.111	-0.206	-0.151	-0.115
8	CORR	LOG(SIZE)	-0.011	0.236	0.038	0.139	1	0.039	0.228	0.149	0.205	-0.093
9	CORR	LIQD	-0.173	0.121	-0.049	-0.409	0.039	1	-0.185	0.11	0.097	0.227
10	CORR	PROPFIX	0.074	-0.125	-0.019	0.111	0.228	-0.185	1	0.165	-0.03	-0.122
11	CORR	P/E	-0.098	0.15	-0.052	-0.206	0.149	0.11	0.165	1	0.277	-0.035
12	CORR	MBV	-0.148	0.27	0.102	-0.151	0.205	0.097	-0.03	0.277	1	0.052
13	CORR	GRDUMMY	0.014	0.042	-0.139	-0.115	-0.093	0.227	-0.122	-0.035	0.052	1

C.3 Correlation Matrix – Industry Relative Data (IRD).

Correlation Matrix - IRD												
Obs	_TYPE_	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY
1	MEAN		0.5	0.619	0.179	0.916	2.059	-2.039	0.853	0.67	0.794	0.276
2	STD		0.502	4.543	2.276	0.785	0.679	43.314	0.342	1.901	0.683	0.448
3	N		156	156	156	156	156	156	156	156	156	156
4	CORR	DV	1	-0.28	-0.205	0.184	0.008	-0.095	0.063	-0.17	-0.173	-0.014
5	CORR	RSHF	-0.28	1	0.407	-0.067	0.104	0.063	-0.033	0.281	0.343	-0.034
6	CORR	TUT	-0.205	0.407	1	-0.015	0.136	-0.129	0.037	0.087	0.243	-0.127
7	CORR	GEAR	0.184	-0.067	-0.015	1	0.224	-0.088	0.155	-0.083	-0.192	-0.002
8	CORR	LOG(SIZE)	0.008	0.104	0.136	0.224	1	-0.06	0.24	0.029	0.205	0.007
9	CORR	LIQ	-0.095	0.063	-0.129	-0.088	-0.06	1	0.199	-0.035	0.025	0.068
10	CORR	PROPFIX	0.063	-0.033	0.037	0.155	0.24	0.199	1	0.099	0.003	0.03
11	CORR	P/E	-0.17	0.281	0.087	-0.083	0.029	-0.035	0.099	1	0.106	-0.096
12	CORR	MBV	-0.173	0.343	0.243	-0.192	0.205	0.025	0.003	0.106	1	0.047
13	CORR	GRDUMMY	-0.014	-0.034	-0.127	-0.002	0.007	0.068	0.03	-0.096	0.047	1

C.4 Correlation Matrix – Time Data (-2).

Correlation Matrix - TD(-2)												
Obs	_TYPE_	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY
1	MEAN		0.5	21.287	8.269	67.114	4.354	0.971	0.426	11.967	12.049	0.321
2	STD		0.502	27.481	23.464	69.171	0.667	0.186	0.641	18.818	11.489	0.468
3	N		156	156	156	156	156	156	156	156	156	156
4	CORR	DV	1	-0.117	-0.045	0.029	0.076	-0.143	-0.008	0.109	-0.087	-0.055
5	CORR	RSHF	-0.117	1	0.371	0.138	0.092	-0.008	-0.185	-0.096	0.227	-0.005
6	CORR	TUT	-0.045	0.371	1	0.087	-0.064	-0.191	-0.167	-0.043	0.125	-0.035
7	CORR	GEAR	0.029	0.138	0.087	1	0.225	-0.32	0.092	-0.004	-0.027	0.02
8	CORR	LOG(SIZE)	0.076	0.092	-0.064	0.225	1	0.02	0.226	-0.162	0.113	-0.114
9	CORR	LIQD	-0.143	-0.008	-0.191	-0.32	0.02	1	-0.074	0.047	0.036	0.093
10	CORR	PROPFIX	-0.008	-0.185	-0.167	0.092	0.226	-0.074	1	0.001	-0.086	0.096
11	CORR	P/E	0.109	-0.096	-0.043	-0.004	-0.162	0.047	0.001	1	0.151	0.042
12	CORR	MBV	-0.087	0.227	0.125	-0.027	0.113	0.036	-0.086	0.151	1	0.082
13	CORR	GRDUMMY	-0.055	-0.005	-0.035	0.02	-0.114	0.093	0.096	0.042	0.082	1

C.5 Correlation Matrix – Variation Data (VD).

Correlation Matrix - VD												
Obs	_TYPE_	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY
1	MEAN		0.5	-1.289	5.585	0.773	0.467	0.077	0.002	-0.412	0.168	-0.006
2	STD		0.502	10.133	22.421	4.301	5.743	0.541	0.23	3.716	1.036	0.551
3	N		156	156	156	156	156	156	156	156	156	156
4	CORR	DV	1	-0.125	-0.267	0.21	-0.278	-0.078	0.126	-0.168	-0.082	0.058
5	CORR	RSHF	-0.125	1	0.298	-0.08	0.153	-0.026	-0.022	0.105	0.172	-0.049
6	CORR	TUT	-0.267	0.298	1	-0.03	0.364	-0.092	-0.092	0.054	0.112	-0.207
7	CORR	GEAR	0.21	-0.08	-0.03	1	-0.15	-0.376	0.111	-0.106	-0.155	0.01
8	CORR	LOG(SIZE)	-0.278	0.153	0.364	-0.15	1	0.407	-0.136	0.254	0.19	-0.147
9	CORR	LIQD	-0.078	-0.026	-0.092	-0.376	0.407	1	-0.238	0.115	0.021	0.048
10	CORR	PROPFIX	0.126	-0.022	-0.092	0.111	-0.136	-0.238	1	0	-0.043	0.056
11	CORR	P/E	-0.168	0.105	0.054	-0.106	0.254	0.115	0	1	0.411	-0.024
12	CORR	MBV	-0.082	0.172	0.112	-0.155	0.19	0.021	-0.043	0.411	1	-0.162
13	CORR	GRDUMMY	0.058	-0.049	-0.207	0.01	-0.147	0.048	0.056	-0.024	-0.162	1

C.6 Correlation Matrix - Mixed Data (MD).

Correlation Matrix - MD										
Obs	_TYPE_	_NAME_	DV	SepChair	NEXOutDir	NEX	XOutDir	DirShHo	BigShHo
1	MEAN		0.5	0.644	4.558	2.827	1.452	12.351	28.979
2	STD		0.502	0.481	5.057	1.903	2.755	15.396	20.051
3	N		104	104	104	104	104	104	104
4	CORR	DV	1	-0.261	-0.048	-0.061	-0.126	-0.133	0.043
5	CORR	SepChair	-0.261	1	0.123	0.139	0.138	-0.035	0.028
6	CORR	NEXOutDir	-0.048	0.123	1	0.659	0.13	-0.463	0.105
7	CORR	NEX	-0.061	0.139	0.659	1	0.188	-0.402	0.068
8	CORR	XOutDir	-0.126	0.138	0.13	0.188	1	-0.19	-0.112
9	CORR	DirShHo	-0.133	-0.035	-0.463	-0.402	-0.19	1	-0.277
10	CORR	BigShHo	0.043	0.028	0.105	0.068	-0.112	-0.277	1

C.7 Correlation Matrix -Time Data (-1).

Correlation Matrix - TD(-1)									
Obs	_TYPE_	_NAME_	DV	SepChair	NEXOutDir	NEX	XOutDir	DirShHo	BigShHo
1	MEAN		0.5	0.603	3.744	2.5	1.397	13.135	27.047
2	STD		0.502	0.491	4.854	1.732	2.333	16.196	19.021
3	N		156	156	156	156	156	156	156
4	CORR	DV	1	-0.236	-0.083	-0.082	0.034	-0.021	0.157
5	CORR	SepChair	-0.236	1	0.205	0.249	0.061	0.01	-0.005
6	CORR	NEXOutDir	-0.083	0.205	1	0.666	0.117	-0.344	0.005
7	CORR	NEX	-0.082	0.249	0.666	1	0.148	-0.373	-0.03
8	CORR	XOutDir	0.034	0.061	0.117	0.148	1	-0.1	-0.156
9	CORR	DirShHo	-0.021	0.01	-0.344	-0.373	-0.1	1	-0.232
10	CORR	BigShHo	0.157	-0.005	0.005	-0.03	-0.156	-0.232	1

C.7 Correlation Matrix -Time Data (-2).

Correlation Matrix - TD(-2)										
Obs	_TYPE_	_NAME_	DV	SepChair	NEXOutDir	NEX	XOutDir	DirShHo	BigShHo
1	MEAN		0.5	0.628	4.083	2.603	1.538	15.338	22.825
2	STD		0.502	0.485	4.977	2.078	2.307	17.679	18.005
3	N		156	156	156	156	156	156	156
4	CORR	DV	1	-0.106	0.012	-0.037	0.002	0.009	0.053
5	CORR	SepChair	-0.106	1	0.114	0.242	0.122	-0.152	0.091
6	CORR	NEXOutDir	0.012	0.114	1	0.725	0.107	-0.512	0.025
7	CORR	NEX	-0.037	0.242	0.725	1	0.13	-0.521	0.1
8	CORR	XOutDir	0.002	0.122	0.107	0.13	1	-0.056	-0.018
9	CORR	DirShHo	0.009	-0.152	-0.512	-0.521	-0.056	1	-0.174
10	CORR	BigShHo	0.053	0.091	0.025	0.1	-0.018	-0.174	1

C.8 Correlation Matrix – Total Mixed Data (TMD).

Correlation Matrix - TMD

TYPE	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY	SepChair	NEXOutDir	NEX	XOutDir	DirShHo	BigShHo
MEAN		0.5	15.075	7.12	67.605	4.476	0.996	0.381	12.557	12.023	0.327	0.644	4.558	2.827	1.452	12.351	28.979
STD		0.502	41.6	18.327	68.473	0.636	0.232	0.184	38.393	10.477	0.471	0.481	5.057	1.903	2.755	15.396	20.051
N		104	104	104	104	104	104	104	104	104	104	104	104	104	104	104	104
CORR	DV	1	-0.369	-0.325	0.198	0	-0.215	0.035	-0.101	-0.246	-0.041	-0.261	-0.048	-0.061	-0.126	-0.133	0.043
CORR	RSHF	-0.369	1	0.369	-0.062	0.313	0.056	0.016	0.012	0.207	0.011	0.045	0.095	0.1	0.012	-0.18	-0.108
CORR	TUT	-0.325	0.369	1	0.015	0.159	-0.152	0.021	-0.102	-0.01	-0.044	0.064	0.039	0.046	0.112	-0.037	-0.098
CORR	GEAR	0.198	-0.062	0.015	1	0.191	-0.326	0.085	-0.195	-0.014	-0.052	-0.065	0.054	0.093	0.188	-0.116	-0.133
CORR	LOG(SIZE)	0	0.313	0.159	0.191	1	0.028	0.429	0.152	0.145	-0.235	0.044	0.421	0.387	0.144	-0.628	-0.004
CORR	LIQD	-0.215	0.056	-0.152	-0.326	0.028	1	-0.081	-0.004	0.024	0.236	0.07	0.115	0.046	-0.112	0.081	0.071
CORR	PROPFIX	0.035	0.016	0.021	0.085	0.429	-0.081	1	0.1	-0.08	-0.053	-0.16	0.182	0.177	0.07	-0.238	0.098
CORR	P/E	-0.101	0.012	-0.102	-0.195	0.152	-0.004	0.1	1	0.303	-0.073	0.233	0.038	-0.03	-0.005	-0.158	0.126
CORR	MBV	-0.246	0.207	-0.01	-0.014	0.145	0.024	-0.08	0.303	1	0.06	0.23	0.013	0.079	0.102	-0.088	0.005
CORR	GRDUMMY	-0.041	0.011	-0.044	-0.052	-0.235	0.236	-0.053	-0.073	0.06	1	-0.124	-0.179	-0.151	-0.157	0.266	-0.044
CORR	SepChair	-0.261	0.045	0.064	-0.065	0.044	0.07	-0.16	0.233	0.23	-0.124	1	0.123	0.139	0.138	-0.035	0.028
CORR	NEXOutDir	-0.048	0.095	0.039	0.054	0.421	0.115	0.182	0.038	0.013	-0.179	0.123	1	0.659	0.13	-0.463	0.105
CORR	NEX	-0.061	0.1	0.046	0.093	0.387	0.046	0.177	-0.03	0.079	-0.151	0.139	0.659	1	0.188	-0.402	0.068
CORR	XOutDir	-0.126	0.012	0.112	0.188	0.144	-0.112	0.07	-0.005	0.102	-0.157	0.138	0.13	0.188	1	-0.19	-0.112
CORR	DirShHo	-0.133	-0.18	-0.037	-0.116	-0.628	0.081	-0.238	-0.158	-0.088	0.266	-0.035	-0.463	-0.402	-0.19	1	-0.277
CORR	BigShHo	0.043	-0.108	-0.098	-0.133	-0.004	0.071	0.098	0.126	0.005	-0.044	0.028	0.105	0.068	-0.112	-0.277	1

C.9 Correlation Matrix – Total Time Data (TTD).

Correlation Matrix - TTD

TYPE	_NAME_	DV	RSHF	TUT	GEAR	LOG(SIZE)	LIQD	PROPFIX	P/E	MBV	GRDUMMY	SepChair	NEXOutDir	NEX	XOutDir	DirShHo	BigShHo
MEAN		0.5	14.295	5.585	67.947	4.418	1.005	0.371	9.567	11.299	0.314	0.603	3.744	2.5	1.397	13.135	27.047
STD		0.502	36.563	22.421	57.37	0.621	0.229	0.165	33.668	9.997	0.466	0.491	4.854	1.732	2.333	16.196	19.021
N		156	156	156	156	156	156	156	156	156	156	156	156	156	156	156	156
CORR	DV	1	-0.265	-0.267	0.159	-0.011	-0.173	0.074	-0.098	-0.148	0.014	-0.236	-0.083	-0.082	0.034	-0.021	0.157
CORR	RSHF	-0.265	1	0.502	-0.043	0.236	0.121	-0.125	0.15	0.27	0.042	0.094	0.145	0.146	-0.017	-0.031	-0.098
CORR	TUT	-0.267	0.502	1	-0.057	0.038	-0.049	-0.019	-0.052	0.102	-0.139	0.107	0.06	0.072	0.035	0.045	-0.078
CORR	GEAR	0.159	-0.043	-0.057	1	0.139	-0.409	0.111	-0.206	-0.151	-0.115	-0.088	0.1	0.092	0.129	-0.214	-0.004
CORR	LOG(SIZE)	-0.011	0.236	0.038	0.139	1	0.039	0.228	0.149	0.205	-0.093	0.042	0.376	0.381	0.155	-0.447	-0.169
CORR	LIQD	-0.173	0.121	-0.049	-0.409	0.039	1	-0.185	0.11	0.097	0.227	0.037	-0.058	-0.106	0.02	0.06	-0.071
CORR	PROPFIX	0.074	-0.125	-0.019	0.111	0.228	-0.185	1	0.165	-0.03	-0.122	-0.089	0.197	0.158	-0.001	-0.098	-0.025
CORR	P/E	-0.098	0.15	-0.052	-0.206	0.149	0.11	0.165	1	0.277	-0.035	0.168	0.1	0.118	-0.009	-0.047	-0.01
CORR	MBV	-0.148	0.27	0.102	-0.151	0.205	0.097	-0.03	0.277	1	0.052	0.023	0.008	0.019	0.03	-0.075	-0.032
CORR	GRDUMMY	0.014	0.042	-0.139	-0.115	-0.093	0.227	-0.122	-0.035	0.052	1	-0.071	-0.139	-0.143	-0.189	0.12	-0.026
CORR	SepChair	-0.236	0.094	0.107	-0.088	0.042	0.037	-0.089	0.168	0.023	-0.071	1	0.205	0.249	0.061	0.01	-0.005
CORR	NEXOutDir	-0.083	0.145	0.06	0.1	0.376	-0.058	0.197	0.1	0.008	-0.139	0.205	1	0.666	0.117	-0.344	0.005
CORR	NEX	-0.082	0.146	0.072	0.092	0.381	-0.106	0.158	0.118	0.019	-0.143	0.249	0.666	1	0.148	-0.373	-0.03
CORR	XOutDir	0.034	-0.017	0.035	0.129	0.155	0.02	-0.001	-0.009	0.03	-0.189	0.061	0.117	0.148	1	-0.1	-0.156
CORR	DirShHo	-0.021	-0.031	0.045	-0.214	-0.447	0.06	-0.098	-0.047	-0.075	0.12	0.01	-0.344	-0.373	-0.1	1	-0.232
CORR	BigShHo	0.157	-0.098	-0.078	-0.004	-0.169	-0.071	-0.025	-0.01	-0.032	-0.026	-0.005	0.005	-0.03	-0.156	-0.232	1

APPENDIX D

LIST OF COMPANIES AND THEIR STATUS

NAME	STATUS	SIC CODE (FAME)	YEAR ACQ.
BENJAMIN PRIEST	Target	3169	1990
MOLINS PLC	Target	3286	1990
EPICURE IND PLC	Target	3289	1990
RACAL ELEC PLC	Target	3433	1990
ALUMASC GROUP PLC	Target	4270	1990
STRONG & FISHER PLC	Target	4410	1990
HOBSON PUBLISH. PLC	Target	4754	1990
CRYSTALATE HOLD. PLC	Target	4836	1990
BIRMINGHAM MINT PLC	Target	4910	1990
UNILOCK PLC	Target	3284	1990
DAVIES & METCALFE PLC	Target	3289	1990
AMS INDUSTRIES PLC	Target	3454	1990
STC PLC	Target	3441	1990
AQUASCUTUM GROUP PLC	Target	4532	1990
SKETCHLEY PLC	Target	4534	1990
AARONSON BROSS PLC	Target	4620	1990
EDBRO PLC	Target	3283	1991
HAWKER SIDDELEY PLC	Target	3420	1991
FLEXELLO CASTORS PLC	Target	3530	1991
JOHN J. LEES PLC	Target	4214	1991
TOOTAL GROUP PLC	Target	4321	1991
DAKS SIMPSON PLC	Target	4532	1991
RITZ DESIGN GROUP PLC	Target	4533	1991
TACE PLC	Target	3284	1991
STEETLEY PLC	Target	3441	1991
BARDON GROUP PLC	Target	2310	1991
CARBO PLC	Target	2460	1991
ROCKWARE GROUP PLC	Target	2470	1991
KINGSGRANGE PLC	Target	2581	1991
THURGAR BARDEX PLC	Target	3142	1991
YALE AND VALOR PLC	Target	3165	1991
DAVY CORPORATION PLC	Target	3240	1991
HERRBURGER BROOKS PLC	Target	4920	1991

GORING KERR PLC	Target	3443	1991
MANDERS HOLD. PLC	Target	2551	1992
MACARTHY PLC	Target	2570	1992
JAMES WILKES PLC	Target	3276	1992
PENNY & GILES INT'L PLC	Target	3442	1992
THOMAS ROBINSON PLC	Target	4160	1992
MORLAND & CO PLC	Target	4283	1992
STAG FURNITURE GROUP PLC	Target	4671	1992
TAVENERS PLC	Target	4214	1992
USHER-WALKER PLC	Target	2552	1992
CONTINUOUS STATIONARY PLC	Target	4723	1992
CRONITE GROUP PLC	Target	3110	1992
WORCESTER GROUP PLC	Target	3165	1992
BRITISH BUILDING ENG. APP. PLC	Target	3289	1992
AMSTRAD PLC	Target	3454	1992
RANKS HOVIS MCDOUGALL PLC	Target	4160	1992
ARTHUR LEE & SONS PLC	Target	2210	1993
MULTITONE ELEC. PLC	Target	3441	1993
HUNTER SAPHIR PLC	Target	4122	1993
HOUSE OF LEROSE PLC	Target	4533	1993
NU-SWIFT PLC	Target	2567	1993
GOODHEAD GROUP PLC	Target	4751	1993
CLIFFORD FOODS PLC	Target	4130	1993
HARRISON IND PLC	Target	3140	1993
FERRANTI INT'L PLC	Target	3204	1993
EVODE GROUP PLC	Target	2562	1993
WHEWAY PLC	Target	3251	1993
INVERGORDON DISTILLERS PLC	Target	3244	1993
BRITISH SYPHON IND PLC	Target	4710	1993
ATTWOODS PLC	Target	2310	1994
LINREAD PLC	Target	3169	1994
MOLYNX HOLDINGS PLC	Target	3222	1994
SCHOLES GROUP PLC	Target	3444	1994
DALE ELEC INT'L PLC	Target	3454	1994
WESTLAND GROUP PLC	Target	3640	1994
TOWLES PLC	Target	4363	1994
MAGELLAN IND. PLC	Target	4532	1994
LEC REFRIGERATION PLC	Target	3284	1994
KEMBREY PLC	Target	3435	1994

CONTROL TECHNIQUES PLC	Target	3454	1994
ELSWICK PLC	Target	3634	1994
VIVAT HOLDING PLC	Target	4534	1994
NMC GROUP PLC	Target	4710	1994
J.W. SPEAR & SONS PLC	Target	4941	1994
DAVENPORT VERNON PLC	Target	4959	1994
C I GROUP PLC	Target	2235	1995
ALFRED MCALPINE PLC	Target	2396	1995
DOBSON PARK INDUSTRIES PLC	Target	2437	1995
VICTAULIC PLC	Target	2515	1995
KALON GROUP PLC	Target	2551	1995
FISONS PLC	Target	2570	1995
WELCOME PLC	Target	2570	1995
SCANTRONIC HOLDINGS PLC	Target	3302	1995
FAIRLINE BOATS PLC	Target	3610	1995
VSEL PLC	Target	3610	1995
DALEPAK FOODS PLC	Target	4122	1995
ATKINS GROUP	Target	4363	1995
CASKET PLC	Target	4557	1995
FINE DECOR PLC	Target	4721	1995
MAGNOLIA GROUP PLC	Target	4728	1995
HARTONS GROUP PLC	Target	2514	1996
AUTOMATED SECURITY (HOLDINGS) PLC	Target	3433	1996
UNITECH PLC	Target	3442	1996
ERF (HOLDINGS) PLC	Target	3510	1996
NEOTRONICS TCHNOLOGY PLC	Target	3710	1996
EVEREST FOODS PLC	Target	4122	1996
MACALLAN-GLENLIVET PLC	Target	4240	1996
THE TELEGRAPH PLC	Target	4751	1996
FERRY PICKERING GROUP PLC	Target	4752	1996
ASPREY PLC.	Target	4910	1996
600 GROUP PLC	Matched	3222	-
A F BULGIN & CO PLC	Matched	3444	-
A.B. ELEC GROUP PLC	Matched	3454	-
A.G. BARR PLC	Matched	4283	-
ABBEY PANELS INV PLC	Matched	4910	-
ABBEYCREST PLC	Matched	4910	-
ADWEST GROUP PLC	Matched	3640	-
AEROSPACE ENGINEERING PLC	Matched	3640	-

AIM GROUP PLC	Matched	3640	-
AIRSPRUNG FURNITURE GROUP PLC	Matched	4671	-
ALEXANDER RUSSELL PLC	Matched	2310	-
ALLIED-LYONS PLC	Matched	4270	-
AMBER DAY HOLD PLC	Matched	4533	-
APV PLC	Matched	3205	-
ARMOUR TRUST PLC	Matched	3454	-
ASH & LACY PLC	Matched	2210	-
ASTEC PLC	Matched	3302	-
ASW HOLDINGS PLC	Matched	2210	-
AUSTIN REED GROUP PLC	Matched	4532	-
AVON RUBBER PLC	Matched	4836	-
BARRY WEHMILLER INT'L PLC	Matched	3244	-
BASS PLC	Matched	4270	-
BBA GROUP PLC	Matched	3255	-
BEALES HUNTER PLC	Matched	4533	-
BEMROSE CORPORATION PLC	Matched	4754	-
BENSON GROUP PLC	Matched	3281	-
BENSONS CRISPS PLC	Matched	3164	-
BLAGDEN INDUSTRIES PLC	Matched	3164	-
BLICK PLC	Matched	3441	-
BODYCOTE INT'L PLC	Matched	3169	-
BOSTROM PLC	Matched	3530	-
BPB INDUSTRIES PLC	Matched	4710	-
BRAITHWAITE PLC	Matched	3284	-
BREEDON PLC	Matched	2310	-
BRIDGEND GROUP PLC	Matched	3169	-
BRISTOL EVENING POST PLC	Matched	4751	-
BRITISH POLYTHENE IND PLC	Matched	4836	-
BRITISH STEEL PLC	Matched	2210	-
BULLOUGH PLC	Matched	4672	-
BURNDENE INVESTMENTS PLC	Matched	3523	-
BURTONWOOD BREWERY PLC	Matched	4270	-
CADBURY SCHWEPPES PLC	Matched	4214	-
CALDERBURN PLC	Matched	4671	-
CASTINGS PLC	Matched	3111	-
CATHAY INTERNATIONAL HOLDINGS PLC	Matched	4671	-
CHAMBERLIN AND HILL PLC	Matched	3111	-
CONRAD CONTINENTAL PLC	Matched	4533	-

CORNWELL PARKER PLC	Matched	4671	-
COSALT PLC	Matched	3523	-
CRADLEY GROUP PLC	Matched	4754	-
CRAIG & ROSE PLC	Matched	2551	-
CRAY ELEC. HOLD PLC	Matched	3454	-
CRESTON LAND & EST PLC	Matched	3142	-
CRITCHLEY GROUP PLC	Matched	4836	-
DAVID S. SMITH (HOLDINGS) PLC	Matched	4640	-
DE LA RUE PLC	Matched	3276	-
DEWHIRST GROUP PLC	Matched	4533	-
DEWHURST PLC	Matched	3442	-
DFS FURNITURE COMPANY PLC	Matched	4671	-
DOLPHIN PACKAGING PLC	Matched	4836	-
DRS DATA AND RESEARCH SERV PLC	Matched	3302	-
ELBIEF PLC	Matched	4650	-
EMAP PLC	Matched	4751	-
EPWIN GROUP PLC	Matched	4834	-
EUROCOPY PLC	Matched	3441	-
EXCALIBUR GROUP PLC	Matched	4910	-
FAREPAK PLC	Matched	4122	-
FARNELL ELEC. PLC	Matched	3454	-
FENNER PLC	Matched	3261	-
FIFE INDMAR PLC	Matched	3120	-
FIRST TECHNOLOGY PLC	Matched	3433	-
FKI PLC	Matched	3420	-
FOLKES GROUP PLC	Matched	4671	-
FORMINSTER PLC	Matched	4533	-
FREDERICK COOPER PLC	Matched	3169	-
G.R. HOLDINGS PLC	Matched	4410	-
GAMES WORKSHOP GROUP PLC	Matched	3169	-
GARTON ENG PLC	Matched	3120	-
GEI INT'L PLC	Matched	3222	-
GENERAL ELECTRIC COMPANY PLC	Matched	3444	-
GOODWIN PLC	Matched	2481	-
GREENE KING PLC	Matched	4270	-
HADEN MACLELLAN PLC	Matched	3230	-
HAMPSON INDS. PLC	Matched	2245	-
HAVELOCK EUROPA PLC	Matched	4672	-
HAZLEWOOD FOODS PLS	Matched	4122	-

HEADWAY PLC	Matched	4671	-
HELENE PLC	Matched	4532	-
HENLYS GROUP PLC	Matched	3510	-
HEWITT GROUP PLC	Matched	2489	-
HICKING PENTECOST PLC	Matched	4363	-
HILL & SMITH HOLD PLC	Matched	3120	-
HOME COUNTIES NEWSPAPERS HOLD PLC	Matched	4751	-
HORACE SMALL APPAREL PLC	Matched	4532	-
HUNTING PLC	Matched	2551	-
JACQUES VERT PLC	Matched	4533	-
JAMES CROPPER PLC	Matched	4710	-
JOHNSTON GROUP PLC	Matched	3283	-
JOHNSTON PRESS PLC	Matched	4751	-
KALAMAZOO COMPUTER GROUP PLC	Matched	3302	-
LAIRD GROUP PLC	Matched	3138	-
LAURA ASHLEY HOLD PLC	Matched	4533	-
LESLIE WISE GROUP PLC	Matched	4533	-
LINCAT GR PLC	Matched	3165	-
LINX PRINTING TECH PLC	Matched	3302	-
LIONHEART PLC	Matched	3161	-
LONDON INT'L GROUP PLC	Matched	2570	-
LOUIS NEWMARK PLC	Matched	3442	-
LUCAS IND PLC	Matched	3530	-
MARSHALLS PLC	Matched	3283	-
MS INT'L PLC	Matched	3111	-
MTM PLC	Matched	2567	-
NEEPSEND PLC	Matched	3222	-
NOBO GROUP PLC	Matched	3166	-
NORMAN HAY PLC	Matched	3138	-
NORTHERN FOODS PLC	Matched	4130	-
OCEONICS GROUP PLC	Matched	3441	-
PATERSON PLC	Matched	2570	-
PEX PLC	Matched	4363	-
PITTARDS PLC	Matched	4410	-
PLATIGNUMN PLC	Matched	4954	-
PLYSU PLC	Matched	4835	-
POLYPIPE PLC	Matched	4834	-
PRESSAC HOLDINGS PLC	Matched	3444	-

RADAMEC GROUP PLC	Matched	3302	-
RADIANT METAL FINISHING PLC	Matched	3138	-
RADSTONE TECH PLC	Matched	3302	-
RAP GROUP PLC	Matched	4812	-
REAL TIME CONTROL PLC	Matched	3302	-
RECORD HOLDINGS PLC	Matched	3161	-
RELYON GROUP PLC	Matched	4671	-
RICHARDS PLC	Matched	4532	-
RM PLC	Matched	3302	-
ROBERT H LOWE PLC	Matched	4532	-
RPC GROUP PLC	Matched	4836	-
SECURITY SERVICES PLC	Matched	3169	-
SHILOH PLC	Matched	2570	-
SIDNEY C BANKS PLC	Matched	4160	-
SILENTNIGHT HOLDINGS PLC	Matched	4671	-
SIMON ENG PLC	Matched	4160	-
SMITHS INDUSTRIES PLC	Matched	3286	-
SOMIC PLC	Matched	2600	-
SPRING RAM CORP PLC	Matched	2489	-
ST IVES PLC	Matched	3276	-
STAVELEY IND PLC	Matched	3111	-
STERLING INDUSTRIES PLC	Matched	3222	-
SYMONDS ENGINEERING PLC	Matched	3169	-
T & N PLC	Matched	2220	-
T CLARKE PLC	Matched	3420	-
TADPOLE TECH PLC	Matched	3302	-
TELEMETRIX PLC	Matched	3302	-
THE GREENALLS GROUP PLC	Matched	4270	-
THE MAYFLOWER CORP PLC	Matched	3510	-
THE SAGE GROUP PLC	Matched	3302	-
THOMAS WALKER PLC	Matched	3169	-
THORNTONS PLC	Matched	4214	-
TOMKINS PLC	Matched	3453	-
TOYE & COMPANY PLC	Matched	3740	-
TRAFFICMASTER PLC	Matched	3302	-
TRANSTEC PLC	Matched	3169	-
TRINITY HOLDINGS PLC	Matched	4811	-
TRIPLEX LLOYD PLC	Matched	3111	-
TUNSTALL GROUP PLC	Matched	3433	-

TURNPYKE GROUP PLC	Matched	3137	-
UNIGATE PLC	Matched	4122	-
UNITED BISCUITS HOLD PLC	Matched	4122	-
UNITED INDUSTRIES PLC	Matched	3222	-
VICKERS PLC	Matched	3276	-
VOLEX GROUP PLC	Matched	3442	-
WAGON INDUSTRIAL HOLD PLC	Matched	3246	-
WALKER & STAFF HOLD PLC	Matched	3283	-
WARDLE STOREYS PLC	Matched	4812	-
WATMOUGHS PLC-M	Matched	4752	-
WESCOL GROUP PLC	Matched	3142	-
WIDNEY PLC	Matched	3444	-
WILLIAM BAIRD PLC	Matched	4532	-
WILLIAM RANSOM & SON PLC	Matched	2570	-
WOLSELEY PLC	Matched	4834	-
WORTHINGTON GROUP PLC	Matched	4321	-
YORKSHIRE CHEMICALS PLC	Matched	2516	-
YULE CATTO & CO PLC	Matched	2567	-
ALBRIGHTON PLC	Matched	2310	-
BRIDGEND	Matched	3169	-
BROWN (DAVID) GROUP PLC	Matched	3222	-
BOWTHORPE PLC	Matched	3444	-
M.L. HOLDING PLC	Matched	3454	-
AEROSPACE ENG. PLC	Matched	3640	-
DAVENPORT KNITWEAR PLC	Matched	4363	-
GENT (S.R.) PLC	Matched	4533	-
BURNFIELD PLC	Matched	3284	-
FORWARD TECH. IND. PLC	Matched	3435	-
VCI PLC	Matched	3452	-
CREST PACKAGING PLC	Matched	4725	-
ALEXANDRA WORKWEAR PLC	Matched	4534	-
BUNZL PLC	Matched	4710	-
BLUEBIRD TOYS PLC	Matched	4941	-
BERISFORD INT. PLC	Matched	4239	-

APPENDIX E

DESCRIPTION AND THE SOURCE OF THE VARIABLES USED

VARIABLE	DEFINITION	SOURCE
Return on Shareholders' Fund (RSHF)	Profit before tax / shareholders funds	FAME
Turnover Trend (TUT)	Turnover (t0) – Turnover (t-1) / Turnover (t-1)	FAME
Gearing (Gear)	(Long term liabilities + Bank overdrafts) / (share capital + reserves)	FAME
Liquidity	(Current Assets – Stocks / Current Liabilities)	FAME
Log(Size)	Log((Fixed Assets + Current Assets) – Current Liabilities)	FAME
Proportional Fixed Assets (Propfix)	Fixed Assets / (Fixed Assets + Current Assets)	FAME
Price Earning Ratio (P/E)	Share Price / Earnings	FT/FAME Datastream
Market to Book Value (MBV)	Market Value / Book Value	FT/FAME Datastream
SepChair	Whether the position of chairman and chief executive are separated	P.W.C.R. A.R.
NEX:	The proportion of outside directors on the board	P.W.C.R. A.R.
NEXOutDir	The weighted average of outside directorships held by the non-executive directors	P.W.C.R. A.R.
XOutDir	The weighted average of outside directorships held by the executive directors	P.W.C.R. A.R.
DirShHo	Percentage of the companies' ordinary shares held by the executive directors	P.W.C.R. A.R.
BigShHo	Outside shareholding of major shareholders	P.W.C.R. A.R.

Notes:

1. FAME: Financial Analysis Made Easy
2. FT: Financial Times.
3. P.W.C.R.: Price Waterhouse Corporate Register
4. A.R.: Annual Reports of the respected Companies

APPENDIX F

LOGISTIC REGRESSION MODEL RESULTS

Time Data (-1) - Financial

The LOGISTIC Procedure

Model Information

Data Set	FIN.TD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	146	186.7436	1.2791	0.0128
Pearson	146	156.4718	1.0717	0.2619

Number of unique profiles: 156

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	206.744
SC	221.312	237.242
-2 Log L	216.262	186.744

R-Square 0.1724 Max-rescaled R-Square 0.2299

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	29.5184	9	0.0005
Score	26.1999	9	0.0019
Wald	20.9463	9	0.0129

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	2.8456	1.6400	3.0107	0.0827
RSHF	1	-0.00968	0.00764	1.6074	0.2049
TUT	1	-0.0254	0.0107	5.7052	0.0169
GEAR	1	0.00260	0.00364	0.5118	0.4744
LOG(SIZE)	1	-0.0781	0.3129	0.0623	0.8030
LIQD	1	-2.2717	0.9749	5.4298	0.0198
PROPFIX	1	0.3577	1.1582	0.0954	0.7574
P/E	1	-0.00826	0.00634	1.6984	0.1925
MBV	1	-0.0188	0.0188	0.9941	0.3187
GRDUMMY	1	0.1549	0.3974	0.1520	0.6966

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.990	0.976	1.005
TUT	0.975	0.955	0.995
GEAR	1.003	0.995	1.010
LOG(SIZE)	0.925	0.501	1.708
LIQD	0.103	0.015	0.697
PROPFIX	1.430	0.148	13.843
P/E	0.992	0.980	1.004
MBV	0.981	0.946	1.018
GRDUMMY	1.168	0.536	2.544

Association of Predicted Probabilities and Observed Responses

Percent Concordant	73.9	Somers' D	0.482
Percent Discordant	25.8	Gamma	0.483
Percent Tied	0.3	Tau-a	0.242
Pairs	6084	c	0.741

Profile Likelihood Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	2.8456	-0.3080	6.1619
RSHF	-0.00968	-0.0277	0.00269
TUT	-0.0254	-0.0479	-0.00620
GEAR	0.00260	-0.00448	0.00991
LOG(SIZE)	-0.0781	-0.6960	0.5395
LIQD	-2.2717	-4.3181	-0.4573
PROPFIX	0.3577	-1.9323	2.6401

P/E	-0.00826	-0.0240	0.00275
MBV	-0.0188	-0.0567	0.0180
GRDUMMY	0.1549	-0.6247	0.9407

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.973	1.003
TUT	1.0000	0.975	0.953	0.994
GEAR	1.0000	1.003	0.996	1.010
LOG(SIZE)	1.0000	0.925	0.499	1.715
LIQD	1.0000	0.103	0.013	0.633
PROPFIX	1.0000	1.430	0.145	14.015
P/E	1.0000	0.992	0.976	1.003
MBV	1.0000	0.981	0.945	1.018
GRDUMMY	1.0000	1.168	0.535	2.562

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.976	1.005
TUT	1.0000	0.975	0.955	0.995
GEAR	1.0000	1.003	0.995	1.010
LOG(SIZE)	1.0000	0.925	0.501	1.708
LIQD	1.0000	0.103	0.015	0.697
PROPFIX	1.0000	1.430	0.148	13.843
P/E	1.0000	0.992	0.980	1.004
MBV	1.0000	0.981	0.946	1.018
GRDUMMY	1.0000	1.168	0.536	2.544

The LOGISTIC Procedure

Model Information

Data Set	FIN.TD_DER2
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	146	200.5604	1.3737	0.0019
Pearson	146	155.2478	1.0633	0.2847

Number of unique profiles: 156

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	220.560
SC	221.312	251.059
-2 Log L	216.262	200.560

R-Square 0.0958 Max-rescaled R-Square 0.1277

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.7016	9	0.0734
Score	13.7972	9	0.1297
Wald	11.9773	9	0.2146

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	2.3468	1.7482	1.8020	0.1795
RSHF	1	-0.0104	0.00781	1.7647	0.1840
TUT	1	-0.00752	0.00907	0.6877	0.4070
GEAR	1	0.000628	0.00308	0.0415	0.8386
LOG(SIZE)	1	0.3800	0.3154	1.4522	0.2282
LIQD	1	-3.2117	1.1903	7.2807	0.0070
PROPFIX	1	-1.2733	1.1148	1.3046	0.2534
P/E	1	0.00176	0.00924	0.0361	0.8493
MBV	1	-0.0122	0.0154	0.6240	0.4296
GRDUMMY	1	-0.1147	0.3722	0.0949	0.7580

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.990	0.975	1.005
TUT	0.993	0.975	1.010
GEAR	1.001	0.995	1.007
LOG(SIZE)	1.462	0.788	2.713
LIQD	0.040	0.004	0.415
PROPFIX	0.280	0.031	2.488
P/E	1.002	0.984	1.020
MBV	0.988	0.959	1.018
GRDUMMY	0.892	0.430	1.849

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.2	Somers' D	0.327
Percent Discordant	33.4	Gamma	0.328
Percent Tied	0.4	Tau-a	0.165
Pairs	6084	c	0.664

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	2.3468	-0.9831	5.9190
RSHF	-0.0104	-0.0269	0.00416
TUT	-0.00752	-0.0257	0.0102
GEAR	0.000628	-0.00562	0.00672
LOG(SIZE)	0.3800	-0.2322	1.0093
LIQD	-3.2117	-5.7352	-1.0347
PROPFIX	-1.2733	-3.5057	0.1940
P/E	0.00176	-0.0172	0.0205
MBV	-0.0122	-0.0433	0.0177
GRDUMMY	-0.1147	-0.8472	0.6183

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.973	1.004
TUT	1.0000	0.993	0.975	1.010
GEAR	1.0000	1.001	0.994	1.007
LOG(SIZE)	1.0000	1.462	0.793	2.744
LIQD	1.0000	0.040	0.003	0.355
PROPFIX	1.0000	0.280	0.030	1.214
P/E	1.0000	1.002	0.983	1.021
MBV	1.0000	0.988	0.958	1.018
GRDUMMY	1.0000	0.892	0.429	1.856

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.975	1.005
TUT	1.0000	0.993	0.975	1.010
GEAR	1.0000	1.001	0.995	1.007
LOG(SIZE)	1.0000	1.462	0.788	2.713
LIQD	1.0000	0.040	0.004	0.415
PROPFIX	1.0000	0.280	0.031	2.488
P/E	1.0000	1.002	0.984	1.020
MBV	1.0000	0.988	0.959	1.018
GRDUMMY	1.0000	0.892	0.430	1.849

Mixed Data (-1) - Financial

The LOGISTIC Procedure

Model Information

Data Set	FIN.MD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	104
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	52
2	0	52

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	94	101.9992	1.0851	0.2690
Pearson	94	96.0421	1.0217	0.4221

Number of unique profiles: 104

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	121.999
SC	148.819	148.443
-2 Log L	144.175	101.999

R-Square 0.3334 Max-rescaled R-Square 0.4445

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	42.1754	9	<.0001
Score	31.6463	9	0.0002
Wald	20.7838	9	0.0136

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.5630	2.4876	2.0516	0.1520
RSHF	1	-0.0459	0.0196	5.4740	0.0193
TUT	1	-0.0518	0.0188	7.6459	0.0057
GEAR	1	0.00917	0.00476	3.7066	0.0542
LOG(SIZE)	1	0.2250	0.4855	0.2148	0.6430
LIQD	1	-3.2318	1.3283	5.9192	0.0150
PROPFIX	1	-0.2918	1.4548	0.0402	0.8410
P/E	1	-0.00880	0.00748	1.3836	0.2395
MBV	1	-0.0467	0.0257	3.3089	0.0689
GRDUMMY	1	0.1751	0.5756	0.0925	0.7610

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.955	0.919	0.993
TUT	0.949	0.915	0.985
GEAR	1.009	1.000	1.019
LOG(SIZE)	1.252	0.484	3.243
LIQD	0.039	0.003	0.534
PROPFIX	0.747	0.043	12.929
P/E	0.991	0.977	1.006
MBV	0.954	0.908	1.004
GRDUMMY	1.191	0.386	3.681

Association of Predicted Probabilities and Observed Responses

Percent Concordant	83.8	Somers' D	0.678
Percent Discordant	16.1	Gamma	0.679
Percent Tied	0.1	Tau-a	0.342
Pairs	2704	c	0.839

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	3.5630	-1.2121	8.6561
RSHF	-0.0459	-0.0878	-0.0112
TUT	-0.0518	-0.0920	-0.0181
GEAR	0.00917	0.000218	0.0190
LOG(SIZE)	0.2250	-0.7239	1.2010
LIQD	-3.2318	-6.1051	-0.7863
PROPFIX	-0.2918	-3.2890	2.6175
P/E	-0.00880	-0.0289	0.00403
MBV	-0.0467	-0.1003	0.00229
GRDUMMY	0.1751	-0.9625	1.3162

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.955	0.916	0.989
TUT	1.0000	0.949	0.912	0.982
GEAR	1.0000	1.009	1.000	1.019
LOG(SIZE)	1.0000	1.252	0.485	3.324
LIQD	1.0000	0.039	0.002	0.456
PROPFIX	1.0000	0.747	0.037	13.701
P/E	1.0000	0.991	0.972	1.004
MBV	1.0000	0.954	0.905	1.002
GRDUMMY	1.0000	1.191	0.382	3.729

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.955	0.919	0.993
TUT	1.0000	0.949	0.915	0.985
GEAR	1.0000	1.009	1.000	1.019
LOG(SIZE)	1.0000	1.252	0.484	3.243
LIQD	1.0000	0.039	0.003	0.534
PROPFIX	1.0000	0.747	0.043	12.929
P/E	1.0000	0.991	0.977	1.006
MBV	1.0000	0.954	0.908	1.004
GRDUMMY	1.0000	1.191	0.386	3.681

Industry Relative Data - Financial

The LOGISTIC Procedure

Model Information

Data Set	FIN.IRD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	130	194.6245	1.4971	0.0002
Pearson	130	152.7021	1.1746	0.0847

Number of unique profiles: 140

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	214.625
SC	221.312	245.123
-2 Log L	216.262	194.625

R-Square 0.1295 Max-rescaled R-Square 0.1727

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	21.6374	9	0.0101
Score	18.3540	9	0.0313
Wald	15.0110	9	0.0906

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.0832	0.6843	0.0148	0.9032
RSHF	1	-0.1512	0.0991	2.3264	0.1272
TUT	1	0.00796	0.0852	0.0087	0.9256
GEAR	1	0.4629	0.2429	3.6334	0.0566
LOG(SIZE)	1	-0.2159	0.2673	0.6523	0.4193
LIQD	1	-0.00382	0.00443	0.7436	0.3885
PROPFIX	1	0.6050	0.5278	1.3139	0.2517
P/E	1	-0.1037	0.1045	0.9844	0.3211
MBV	1	-0.4313	0.2840	2.3070	0.1288
GRDUMMY	1	-0.0626	0.3822	0.0269	0.8698

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.860	0.708	1.044
TUT	1.008	0.853	1.191
GEAR	1.589	0.987	2.557
LOG(SIZE)	0.806	0.477	1.361
LIQD	0.996	0.988	1.005
PROPFIX	1.831	0.651	5.152
P/E	0.902	0.735	1.106
MBV	0.650	0.372	1.133
GRDUMMY	0.939	0.444	1.987

Association of Predicted Probabilities and Observed Responses

Percent Concordant	69.3	Somers' D	0.390
Percent Discordant	30.3	Gamma	0.391
Percent Tied	0.4	Tau-a	0.196
Pairs	6084	c	0.695

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	0.0832	-1.2608	1.4361
RSHF	-0.1512	-0.3616	0.00617
TUT	0.00796	-0.1704	0.1767
GEAR	0.4629	-0.00528	0.9539
LOG(SIZE)	-0.2159	-0.7509	0.3045
LIQD	-0.00382	-0.0127	0.00475
PROPFIX	0.6050	-0.4212	1.6617
P/E	-0.1037	-0.3299	0.0904
MBV	-0.4313	-1.0130	0.1097
GRDUMMY	-0.0626	-0.8173	0.6877

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.860	0.697	1.006
TUT	1.0000	1.008	0.843	1.193
GEAR	1.0000	1.589	0.995	2.596
LOG(SIZE)	1.0000	0.806	0.472	1.356
LIQD	1.0000	0.996	0.987	1.005
PROPFIX	1.0000	1.831	0.656	5.268
P/E	1.0000	0.902	0.719	1.095
MBV	1.0000	0.650	0.363	1.116
GRDUMMY	1.0000	0.939	0.442	1.989

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.860	0.708	1.044
TUT	1.0000	1.008	0.853	1.191
GEAR	1.0000	1.589	0.987	2.557
LOG(SIZE)	1.0000	0.806	0.477	1.361
LIQD	1.0000	0.996	0.988	1.005
PROPFIX	1.0000	1.831	0.651	5.152
P/E	1.0000	0.902	0.735	1.106
MBV	1.0000	0.650	0.372	1.133
GRDUMMY	1.0000	0.939	0.444	1.987

Variation Data-Financial

The LOGISTIC Procedure

Model Information

Data Set	FIN.VD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	146	188.7886	1.2931	0.0098
Pearson	146	151.1913	1.0356	0.3672

Number of unique profiles: 156

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	208.789
SC	221.312	239.287
-2 Log L	216.262	188.789

R-Square 0.1615 Max-rescaled R-Square 0.2153

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	27.4734	9	0.0012
Score	17.8357	9	0.0371
Wald	17.0134	9	0.0485

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.2361	0.1937	1.4864	0.2228
RSHF	1	-0.0746	0.0639	1.3662	0.2425
TUT	1	-0.0250	0.0102	5.9929	0.0144
GEAR	1	0.0181	0.0576	0.0987	0.7534
LOG(SIZE)	1	-9.5456	4.5760	4.3514	0.0370
LIQD	1	-0.7555	0.9321	0.6570	0.4176
PROPFIX	1	0.8329	0.8882	0.8793	0.3484
P/E	1	0.0327	0.0556	0.3466	0.5561
MBV	1	-0.2599	0.1912	1.8478	0.1740
GRDUMMY	1	-0.1203	0.3302	0.1328	0.7156

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.928	0.819	1.052
TUT	0.975	0.956	0.995
GEAR	1.018	0.910	1.140
LOG(SIZE)	<0.001	<0.001	0.562
LIQD	0.470	0.076	2.920
PROPFIX	2.300	0.403	13.114
P/E	1.033	0.927	1.152
MBV	0.771	0.530	1.122
GRDUMMY	0.887	0.464	1.694

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.7	Somers' D	0.457
Percent Discordant	26.9	Gamma	0.459
Percent Tied	0.4	Tau-a	0.230
Pairs	6084	c	0.729

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	0.2361	-0.1393	0.6228
RSHF	-0.0746	-0.2426	-0.00149
TUT	-0.0250	-0.0463	-0.00617
GEAR	0.0181	-0.0742	0.1583
LOG(SIZE)	-9.5456	-19.9679	-1.6749
LIQD	-0.7555	-2.6513	1.0587
PROPFIX	0.8329	-0.8496	2.6849
P/E	0.0327	-0.0652	0.1832
MBV	-0.2599	-0.6892	0.0900
GRDUMMY	-0.1203	-0.7744	0.5299

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.928	0.785	0.999
TUT	1.0000	0.975	0.955	0.994
GEAR	1.0000	1.018	0.929	1.172
LOG(SIZE)	1.0000	<0.001	<0.001	0.187
LIQD	1.0000	0.470	0.071	2.883
PROPFIX	1.0000	2.300	0.428	14.657
P/E	1.0000	1.033	0.937	1.201
MBV	1.0000	0.771	0.502	1.094
GRDUMMY	1.0000	0.887	0.461	1.699

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.928	0.819	1.052
TUT	1.0000	0.975	0.956	0.995
GEAR	1.0000	1.018	0.910	1.140
LOG(SIZE)	1.0000	<0.001	<0.001	0.562
LIQD	1.0000	0.470	0.076	2.920
PROPFIX	1.0000	2.300	0.403	13.114
P/E	1.0000	1.033	0.927	1.152
MBV	1.0000	0.771	0.530	1.122
GRDUMMY	1.0000	0.887	0.464	1.694

Time Data (-1) - Non-Financial

The LOGISTIC Procedure

Model Information

Data Set	NONFIN.TD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	149	201.1166	1.3498	0.0029
Pearson	149	154.7632	1.0387	0.3564

Number of unique profiles: 156

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	215.117
SC	221.312	236.466
-2 Log L	216.262	201.117

R-Square 0.0925 Max-rescaled R-Square 0.1234

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.1453	6	0.0192
Score	14.5297	6	0.0242
Wald	13.4446	6	0.0365

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.1725	0.5313	0.1054	0.7454
SepChair	1	-0.9473	0.3580	7.0002	0.0082
NEXOutDir	1	-0.0595	0.0496	1.4378	0.2305
NEX	1	0.0438	0.1316	0.1109	0.7391
XOutDir	1	-0.00242	0.0750	0.0010	0.9743
DirShHo	1	0.00107	0.0119	0.0080	0.9289
BigShHo	1	0.0185	0.00983	3.5432	0.0598

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SepChair	0.388	0.192	0.782
NEXOutDir	0.942	0.855	1.038
NEX	1.045	0.807	1.352
XOutDir	0.998	0.861	1.155
DirShHo	1.001	0.978	1.025
BigShHo	1.019	0.999	1.038

Association of Predicted Probabilities and Observed Responses

Percent Concordant	67.5	Somers' D	0.352
Percent Discordant	32.3	Gamma	0.353
Percent Tied	0.2	Tau-a	0.177
Pairs	6084	c	0.676

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	0.1725	-0.8728	1.2240
SepChair	-0.9473	-1.6624	-0.2542
NEXOutDir	-0.0595	-0.1623	0.0334
NEX	0.0438	-0.2150	0.3051
XOutDir	-0.00242	-0.1544	0.1447
DirShHo	0.00107	-0.0226	0.0246
BigShHo	0.0185	-0.00044	0.0383

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.388	0.190	0.776
NEXOutDir	1.0000	0.942	0.850	1.034
NEX	1.0000	1.045	0.807	1.357
XOutDir	1.0000	0.998	0.857	1.156
DirShHo	1.0000	1.001	0.978	1.025
BigShHo	1.0000	1.019	1.000	1.039

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.388	0.192	0.782
NEXOutDir	1.0000	0.942	0.855	1.038
NEX	1.0000	1.045	0.807	1.352
XOutDir	1.0000	0.998	0.861	1.155
DirShHo	1.0000	1.001	0.978	1.025
BigShHo	1.0000	1.019	0.999	1.038

Time Data (-2) - Non-Financial

The LOGISTIC Procedure

Model Information

Data Set	NONFIN.TD_DER2
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	146	212.9018	1.4582	0.0003
Pearson	146	155.7907	1.0671	0.2744

Number of unique profiles: 153

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	226.902
SC	221.312	248.251
-2 Log L	216.262	212.902

R-Square 0.0213 Max-rescaled R-Square 0.0284

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.3601	6	0.7625
Score	3.3287	6	0.7666
Wald	3.2673	6	0.7746

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.00641	0.4655	0.0002	0.9890
SepChair	1	-0.4644	0.3534	1.7268	0.1888
NEXOutDir	1	-0.0311	0.0477	0.4260	0.5139
NEX	1	-0.0707	0.1118	0.4001	0.5270
XOutDir	1	-0.0192	0.0715	0.0723	0.7881
DirShHo	1	0.00549	0.0103	0.2842	0.5939
BigShHo	1	0.00761	0.00961	0.6276	0.4282

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SepChair	0.628	0.314	1.256
NEXOutDir	0.969	0.883	1.064
NEX	1.073	0.862	1.336
XOutDir	0.981	0.853	1.128
DirShHo	1.006	0.985	1.026
BigShHo	1.008	0.989	1.027

Association of Predicted Probabilities and Observed Responses

Percent Concordant	56.1	Somers' D	0.131
Percent Discordant	42.9	Gamma	0.133
Percent Tied	1.0	Tau-a	0.066
Pairs	6084	c	0.566

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	0.00641	-0.9164	0.9220
SepChair	-0.4644	-1.1647	0.2253
NEXOutDir	-0.0311	-0.1296	0.0605
NEX	0.0707	-0.1479	0.3043
XOutDir	-0.0192	-0.1649	0.1215
DirShHo	0.00549	-0.0146	0.0261
BigShHo	0.00761	-0.0112	0.0267

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.628	0.312	1.253
NEXOutDir	1.0000	0.969	0.878	1.062
NEX	1.0000	1.073	0.863	1.356
XOutDir	1.0000	0.981	0.848	1.129
DirShHo	1.0000	1.006	0.985	1.026
BigShHo	1.0000	1.008	0.989	1.027

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.628	0.314	1.256
NEXOutDir	1.0000	0.969	0.883	1.064
NEX	1.0000	1.073	0.862	1.336
XOutDir	1.0000	0.981	0.853	1.128
DirShHo	1.0000	1.006	0.985	1.026
BigShHo	1.0000	1.008	0.989	1.027

Mixed Data - Non-Financial

The LOGISTIC Procedure

Model Information

Data Set	NONFIN.MD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	104
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	52
2	0	52

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	97	124.4245	1.2827	0.0317
Pearson	97	103.9876	1.0720	0.2954

Number of unique profiles: 104

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	138.425
SC	148.819	156.935
-2 Log L	144.175	124.425

R-Square 0.1730 Max-rescaled R-Square 0.2306

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	19.7501	6	0.0031
Score	16.0828	6	0.0133
Wald	12.5289	6	0.0512

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	2.3875	0.8323	8.2282	0.0041
SepChair	1	-1.2198	0.4750	6.5956	0.0102
NEXOutDir	1	-0.0427	0.0582	0.5366	0.4639
NEX	1	-0.0825	0.1490	0.3067	0.5797
XOutDir	1	-0.2912	0.1457	3.9982	0.0455
DirShHo	1	-0.0462	0.0192	5.7839	0.0162
BigShHo	1	-0.00898	0.0119	0.5678	0.4511

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SepChair	0.295	0.116	0.749
NEXOutDir	0.958	0.855	1.074
NEX	0.921	0.688	1.233
XOutDir	0.747	0.562	0.994
DirShHo	0.955	0.920	0.991
BigShHo	0.991	0.968	1.014

Association of Predicted Probabilities and Observed Responses

Percent Concordant	75.5	Somers' D	0.512
Percent Discordant	24.3	Gamma	0.513
Percent Tied	0.1	Tau-a	0.258
Pairs	2704	c	0.756

Profile Likelihood Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	2.3875	0.8515	4.1442
SepChair	-1.2198	-2.1939	-0.3166
NEXOutDir	-0.0427	-0.1590	0.0715
NEX	-0.0825	-0.3827	0.2085
XOutDir	-0.2912	-0.6158	-0.0484
DirShHo	-0.0462	-0.0861	-0.0102
BigShHo	-0.00898	-0.0328	0.0143

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.295	0.111	0.729
NEXOutDir	1.0000	0.958	0.853	1.074
NEX	1.0000	0.921	0.682	1.232
XOutDir	1.0000	0.747	0.540	0.953
DirShHo	1.0000	0.955	0.918	0.990
BigShHo	1.0000	0.991	0.968	1.014

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
SepChair	1.0000	0.295	0.116	0.749
NEXOutDir	1.0000	0.958	0.855	1.074
NEX	1.0000	0.921	0.688	1.233
XOutDir	1.0000	0.747	0.562	0.994
DirShHo	1.0000	0.955	0.920	0.991
BigShHo	1.0000	0.991	0.968	1.014

Variation Data - Non-Financial

The LOGISTIC Procedure

Model Information

Data Set	NONFIN.VD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Complete separation of data points detected.

WARNING: The maximum likelihood estimate does not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	144	0.0314	0.0002	1.0000
Pearson	144	0.0157	0.0001	1.0000

Number of unique profiles: 152

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	16.031
SC	221.312	40.430
-2 Log L	216.262	0.031

R-Square 0.7499 Max-rescaled R-Square 0.9999

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	216.2305	7	<.0001
Score	156.0000	7	<.0001
Wald	1.3310	7	0.9876

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-9.2027	11.3569	0.6566	0.4178
DV	1	18.4055	16.3605	1.2656	0.2606
SepChair	1	-398E-18	15.5605	0.0000	1.0000
NEXOutDir	1	9.66E-17	2.5922	0.0000	1.0000
NEX	1	-187E-18	5.2979	0.0000	1.0000
XOutDir	1	3.23E-16	3.5613	0.0000	1.0000
DirShHo	1	2.84E-17	0.7374	0.0000	1.0000
BigShHo	1	-106E-18	0.4942	0.0000	1.0000

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
DV	>999.999	<0.001	>999.999
SepChair	1.000	<0.001	>999.999
NEXOutDir	1.000	0.006	160.868
NEX	1.000	<0.001	>999.999
XOutDir	1.000	<0.001	>999.999
DirShHo	1.000	0.236	4.243
BigShHo	1.000	0.380	2.634

Association of Predicted Probabilities and Observed Responses

Percent Concordant	100.0	Somers' D	1.000
Percent Discordant	0.0	Gamma	1.000
Percent Tied	0.0	Tau-a	0.503
Pairs	6084	c	1.000

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	-9.2027	.	-3.6690
DV	18.4055	8.7085	.
SepChair	-398E-18	.	.
NEXOutDir	9.66E-17	.	.
NEX	-187E-18	.	.
XOutDir	3.23E-16	.	.
DirShHo	2.84E-17	.	.
BigShHo	-106E-18	.	.

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
DV	1.0000	>999.999	>999.999	.
SepChair	1.0000	1.000	.	.
NEXOutDir	1.0000	1.000	.	.
NEX	1.0000	1.000	.	.
XOutDir	1.0000	1.000	.	.
DirShHo	1.0000	1.000	.	.
BigShHo	1.0000	1.000	.	.

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
DV	1.0000	>999.999	<0.001	>999.999
SepChair	1.0000	1.000	<0.001	>999.999
NEXOutDir	1.0000	1.000	0.006	160.868
NEX	1.0000	1.000	<0.001	>999.999
XOutDir	1.0000	1.000	<0.001	>999.999
DirShHo	1.0000	1.000	0.236	4.243
BigShHo	1.0000	1.000	0.380	2.634

Total Time Data (-1) - Combined

The LOGISTIC Procedure

Model Information

Data Set	COMB.TTD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	140	174.5368	1.2467	0.0253
Pearson	140	156.9870	1.1213	0.1548

Number of unique profiles: 156

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	206.537
SC	221.312	255.334
-2 Log L	216.262	174.537

R-Square 0.2347 Max-rescaled R-Square 0.3129

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	41.7252	15	0.0002
Score	36.6608	15	0.0014
Wald	28.7073	15	0.0175

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.7845	1.9955	0.1545	0.6942
RSHF	1	-0.0102	0.00747	1.8723	0.1712
TUT	1	-0.0228	0.0109	4.4182	0.0356
GEAR	1	0.00267	0.00390	0.4688	0.4935
LOG(SIZE)	1	0.4142	0.4012	1.0658	0.3019
LIQD	1	-2.3957	1.0464	5.2418	0.0221
PROPFIX	1	0.2335	1.2360	0.0357	0.8502
P/E	1	-0.00624	0.00683	0.8352	0.3608
MBV	1	-0.0255	0.0201	1.6144	0.2039
GRDUMMY	1	0.1332	0.4233	0.0991	0.7530
SepChair	1	-0.7817	0.3973	3.8722	0.0491
NEXOutDir	1	-0.0509	0.0550	0.8570	0.3546
NEX	1	-0.0239	0.1501	0.0254	0.8733
XOutDir	1	-0.0107	0.0821	0.0170	0.8964
DirShHo	1	0.0139	0.0142	0.9687	0.3250
BigShHo	1	0.0245	0.0112	4.7801	0.0288

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.990	0.975	1.004
TUT	0.977	0.957	0.998
GEAR	1.003	0.995	1.010
LOG(SIZE)	1.513	0.689	3.322
LIQD	0.091	0.012	0.708
PROPFIX	1.263	0.112	14.240
P/E	0.994	0.981	1.007
MBV	0.975	0.937	1.014
GRDUMMY	1.143	0.498	2.619
SepChair	0.458	0.210	0.997
NEXOutDir	0.950	0.853	1.059
NEX	0.976	0.728	1.310
XOutDir	0.989	0.842	1.162
DirShHo	1.014	0.986	1.043
BigShHo	1.025	1.003	1.048

Association of Predicted Probabilities and Observed Responses

Percent Concordant	78.5	Somers' D	0.571
Percent Discordant	21.4	Gamma	0.572
Percent Tied	0.2	Tau-a	0.287
Pairs	6084	c	0.786

Profile Likelihood Confidence Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	0.7845	-3.1497	4.7333
RSHF	-0.0102	-0.0279	0.00214
TUT	-0.0228	-0.0458	-0.00321
GEAR	0.00267	-0.00490	0.0105
LOG(SIZE)	0.4142	-0.3591	1.2261
LIQD	-2.3957	-4.5973	-0.4523
PROPFIX	0.2335	-2.2220	2.6588
P/E	-0.00624	-0.0226	0.00563
MBV	-0.0255	-0.0661	0.0135
GRDUMMY	0.1332	-0.6968	0.9720
SepChair	-0.7817	-1.5742	-0.00972
NEXOutDir	-0.0509	-0.1657	0.0517
NEX	-0.0239	-0.3221	0.2718
XOutDir	-0.0107	-0.1775	0.1499
DirShHo	0.0139	-0.0137	0.0423
BigShHo	0.0245	0.00303	0.0472

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.972	1.002
TUT	1.0000	0.977	0.955	0.997
GEAR	1.0000	1.003	0.995	1.011
LOG(SIZE)	1.0000	1.513	0.698	3.408
LIQD	1.0000	0.091	0.010	0.636
PROPFIX	1.0000	1.263	0.108	14.278
P/E	1.0000	0.994	0.978	1.006
MBV	1.0000	0.975	0.936	1.014
GRDUMMY	1.0000	1.143	0.498	2.643
SepChair	1.0000	0.458	0.207	0.990
NEXOutDir	1.0000	0.950	0.847	1.053
NEX	1.0000	0.976	0.725	1.312
XOutDir	1.0000	0.989	0.837	1.162
DirShHo	1.0000	1.014	0.986	1.043
BigShHo	1.0000	1.025	1.003	1.048

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.990	0.975	1.004
TUT	1.0000	0.977	0.957	0.998
GEAR	1.0000	1.003	0.995	1.010
LOG(SIZE)	1.0000	1.513	0.689	3.322
LIQD	1.0000	0.091	0.012	0.708
PROPFIX	1.0000	1.263	0.112	14.240
P/E	1.0000	0.994	0.981	1.007
MBV	1.0000	0.975	0.937	1.014
GRDUMMY	1.0000	1.143	0.498	2.619
SepChair	1.0000	0.458	0.210	0.997
NEXOutDir	1.0000	0.950	0.853	1.059
NEX	1.0000	0.976	0.728	1.310
XOutDir	1.0000	0.989	0.842	1.162
DirShHo	1.0000	1.014	0.986	1.043
BigShHo	1.0000	1.025	1.003	1.048

Total Reduced Time Data (-1) - Combined

The LOGISTIC Procedure

Model Information

Data Set	COMB.TD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	156
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	78
2	0	78

Probability modeled is DV='1'.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
36.6608	15	0.0014

Step 1. Effect RSHF entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	208.259
SC	221.312	214.358
-2 Log L	216.262	204.259

R-Square 0.0741 Max-rescaled R-Square 0.0987

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12.0032	1	0.0005
Score	9.2871	1	0.0023
Wald	8.1753	1	0.0042

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
27.3407	14	0.0174

Step 2. Effect SepChair entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	202.771
SC	221.312	211.921
-2 Log L	216.262	196.771

R-Square 0.1174 Max-rescaled R-Square 0.1566

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	19.4905	2	<.0001
Score	17.0070	2	0.0002
Wald	14.4637	2	0.0007

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
20.8155	13	0.0767

Step 3. Effect LIQD entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	199.445
SC	221.312	211.645
-2 Log L	216.262	191.445

R-Square 0.1471 Max-rescaled R-Square 0.1961

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	24.8167	3	<.0001
Score	21.7969	3	<.0001
Wald	17.9005	3	0.0005

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
15.9264	12	0.1946

Step 4. Effect TUT entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	196.241
SC	221.312	211.490
-2 Log L	216.262	186.241

R-Square 0.1751 Max-rescaled R-Square 0.2334

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	30.0212	4	<.0001
Score	26.5690	4	<.0001

Wald	21.4464	4	0.0003

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
11.5999	11	0.3945

Step 5. Effect RSHF is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	198.615
SC	221.312	210.814
-2 Log L	216.262	190.615

R-Square 0.1516 Max-rescaled R-Square 0.2021

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	25.6473	3	<.0001
Score	22.9654	3	<.0001
Wald	19.0562	3	0.0003

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
15.7446	12	0.2032

Step 6. Effect P/E entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates

AIC	218.262	196.752
SC	221.312	212.001
-2 Log L	216.262	186.752

R-Square 0.1724 Max-rescaled R-Square 0.2298

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	29.5103	4	<.0001
Score	26.1470	4	<.0001
Wald	21.2224	4	0.0003

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
11.7359	11	0.3838

Step 7. Effect P/E is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	218.262	198.615
SC	221.312	210.814
-2 Log L	216.262	190.615

R-Square 0.1516 Max-rescaled R-Square 0.2021

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	25.6473	3	<.0001
Score	22.9654	3	<.0001
Wald	19.0562	3	0.0003

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
15.7446	12	0.2032

NOTE: Model building terminates because the last effect entered is removed by the Wald statistic criterion.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	RSHF		1	1	9.2871	.	0.0023
2	SepChair		1	2	7.4478	.	0.0064
3	LIQD		1	3	5.1967	.	0.0226
4	TUT		1	4	4.7581	.	0.0292
5		RSHF	1	3	.	2.9473	0.0860
6	P/E		1	4	4.2936	.	0.0383
7		P/E	1	3	.	2.7787	0.0955

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	152	190.6146	1.2540	0.0184
Pearson	152	155.9088	1.0257	0.3973

Number of unique profiles: 156

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.0077	0.9082	10.9677	0.0009
TUT	1	-0.0290	0.0101	8.2629	0.0040
LIQD	1	-2.3399	0.8542	7.5044	0.0062
SepChair	1	-0.8671	0.3548	5.9749	0.0145

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TUT	0.971	0.952	0.991
LIQD	0.096	0.018	0.514
SepChair	0.420	0.210	0.842

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.2	Somers' D	0.447
Percent Discordant	27.5	Gamma	0.448
Percent Tied	0.2	Tau-a	0.225
Pairs	6084	c	0.724

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	3.0077	1.3140	4.8989

TUT	-0.0290	-0.0501	-0.0106
LIQD	-2.3399	-4.1149	-0.7412
SepChair	-0.8671	-1.5738	-0.1787

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
TUT	1.0000	0.971	0.951	0.989
LIQD	1.0000	0.096	0.016	0.477
SepChair	1.0000	0.420	0.207	0.836

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
TUT	1.0000	0.971	0.952	0.991
LIQD	1.0000	0.096	0.018	0.514
SepChair	1.0000	0.420	0.210	0.842

Total Mixed Data - Combined

The LOGISTIC Procedure

Model Information

Data Set	COMB.TMD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	104
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	52
2	0	52

Probability modeled is DV='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	88	83.9706	0.9542	0.6018
Pearson	88	89.1140	1.0127	0.4468

Number of unique profiles: 104

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	115.971
SC	148.819	158.281
-2 Log L	144.175	83.971

R-Square 0.4395 Max-rescaled R-Square 0.5860

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	60.2040	15	<.0001
Score	41.9178	15	0.0002
Wald	24.6341	15	0.0551

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	7.2264	3.6997	3.8151	0.0508
RSHF	1	-0.0570	0.0216	6.9801	0.0082
TUT	1	-0.0506	0.0205	6.0944	0.0136
GEAR	1	0.00970	0.00513	3.5739	0.0587
LOG(SIZE)	1	-0.0168	0.7246	0.0005	0.9815
LIQD	1	-2.1281	1.4595	2.1260	0.1448
PROPFIX	1	-0.9912	1.8567	0.2850	0.5934
P/E	1	-0.00650	0.0104	0.3940	0.5302
MBV	1	-0.0513	0.0292	3.0916	0.0787
GRDUMMY	1	-0.5515	0.6715	0.6745	0.4115
SepChair	1	-1.4627	0.6594	4.9202	0.0265
NEXOutDir	1	-0.0520	0.0785	0.4389	0.5077
NEX	1	-0.1429	0.1952	0.5364	0.4639
XOutDir	1	-0.3766	0.1902	3.9197	0.0477
DirShHo	1	-0.0573	0.0297	3.7107	0.0541
BigShHo	1	-0.00822	0.0166	0.2440	0.6213

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.945	0.905	0.985
TUT	0.951	0.913	0.990
GEAR	1.010	1.000	1.020
LOG(SIZE)	0.983	0.238	4.069
LIQD	0.119	0.007	2.080
PROPFIX	0.371	0.010	14.123
P/E	0.994	0.974	1.014
MBV	0.950	0.897	1.006
GRDUMMY	0.576	0.155	2.148
SepChair	0.232	0.064	0.843
NEXOutDir	0.949	0.814	1.107
NEX	0.867	0.591	1.271
XOutDir	0.686	0.473	0.996
DirShHo	0.944	0.891	1.001
BigShHo	0.992	0.960	1.025

Association of Predicted Probabilities and Observed Responses

Percent Concordant	89.4	Somers' D	0.790
Percent Discordant	10.4	Gamma	0.791
Percent Tied	0.1	Tau-a	0.399
Pairs	2704	c	0.895

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	7.2264	0.1532	14.9280
RSHF	-0.0570	-0.1034	-0.0191

TUT	-0.0506	-0.0948	-0.0137
GEAR	0.00970	0.000506	0.0209
LOG(SIZE)	-0.0168	-1.4608	1.4333
LIQD	-2.1281	-5.1852	0.6450
PROPFIX	-0.9912	-4.5845	2.9922
P/E	-0.00650	-0.0337	0.00970
MBV	-0.0513	-0.1116	0.00517
GRDUMMY	-0.5515	-1.9111	0.7586
SepChair	-1.4627	-2.8608	-0.2280
NEXOutDir	-0.0520	-0.2103	0.1019
NEX	-0.1429	-0.5453	0.2311
XOutDir	-0.3766	-0.8234	-0.0831
DirShHo	-0.0573	-0.1207	-0.00238
BigShHo	-0.00822	-0.0420	0.0241

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.945	0.902	0.981
TUT	1.0000	0.951	0.910	0.986
GEAR	1.0000	1.010	1.001	1.021
LOG(SIZE)	1.0000	0.983	0.232	4.193
LIQD	1.0000	0.119	0.006	1.906
PROPFIX	1.0000	0.371	0.010	19.929
P/E	1.0000	0.994	0.967	1.010
MBV	1.0000	0.950	0.894	1.005
GRDUMMY	1.0000	0.576	0.148	2.135
SepChair	1.0000	0.232	0.057	0.796
NEXOutDir	1.0000	0.949	0.810	1.107
NEX	1.0000	0.867	0.580	1.260
XOutDir	1.0000	0.686	0.439	0.920
DirShHo	1.0000	0.944	0.886	0.998
BigShHo	1.0000	0.992	0.959	1.024

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.945	0.905	0.985
TUT	1.0000	0.951	0.913	0.990
GEAR	1.0000	1.010	1.000	1.020
LOG(SIZE)	1.0000	0.983	0.238	4.069
LIQD	1.0000	0.119	0.007	2.080
PROPFIX	1.0000	0.371	0.010	14.123
P/E	1.0000	0.994	0.974	1.014
MBV	1.0000	0.950	0.897	1.006
GRDUMMY	1.0000	0.576	0.155	2.148
SepChair	1.0000	0.232	0.064	0.843
NEXOutDir	1.0000	0.949	0.814	1.107
NEX	1.0000	0.867	0.591	1.271
XOutDir	1.0000	0.686	0.473	0.996
DirShHo	1.0000	0.944	0.891	1.001
BigShHo	1.0000	0.992	0.960	1.025

Total Reduced Mixed Data - Combined

The LOGISTIC Procedure

Model Information

Data Set	COMB.TMD_DER1
Response Variable	DV
Number of Response Levels	2
Number of Observations	104
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile

Ordered Value	DV	Total Frequency
1	1	52
2	0	52

Probability modeled is DV='1'.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
41.9178	15	0.0002

Step 1. Effect TUT entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	136.671
SC	148.819	141.960
-2 Log L	144.175	132.671

R-Square 0.1047 Max-rescaled R-Square 0.1396

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.5034	1	0.0007
Score	10.4277	1	0.0012
Wald	9.1007	1	0.0026

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
35.0941	14	0.0014

Step 2. Effect SepChair entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	131.833
SC	148.819	139.766
-2 Log L	144.175	125.833

R-Square 0.1617 Max-rescaled R-Square 0.2156

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	18.3415	2	0.0001
Score	16.4246	2	0.0003
Wald	14.0622	2	0.0009

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
29.9110	13	0.0049

Step 3. Effect RSHF entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	120.584
SC	148.819	131.162
-2 Log L	144.175	112.584

R-Square 0.2620 Max-rescaled R-Square 0.3493

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	31.5903	3	<.0001
Score	23.2278	3	<.0001
Wald	17.3092	3	0.0006

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
23.7781	12	0.0218

Step 4. Effect DirShHo entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	115.674
SC	148.819	128.896
-2 Log L	144.175	105.674

R-Square 0.3094 Max-rescaled R-Square 0.4125

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	38.5006	4	<.0001

Score	27.4531	4	<.0001
Wald	20.3197	4	0.0004

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
18.7541	11	0.0657

Step 5. Effect XOutDir entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	110.971
SC	148.819	126.838
-2 Log L	144.175	98.971

R-Square 0.3525 Max-rescaled R-Square 0.4700

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	45.2033	5	<.0001
Score	29.7906	5	<.0001
Wald	20.5713	5	0.0010

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
13.2040	10	0.2125

Step 6. Effect TUT is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	146.175	113.061
SC	148.819	126.283
-2 Log L	144.175	103.061

R-Square 0.3265 Max-rescaled R-Square 0.4354

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	41.1140	4	<.0001
Score	25.7309	4	<.0001
Wald	19.3324	4	0.0007

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
16.4282	11	0.1260

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
1	TUT		1	1	10.4277	.	0.0012
2	SepChair		1	2	6.7314	.	0.0095
3	RSHF		1	3	7.8811	.	0.0050
4	DirShHo		1	4	6.3768	.	0.0116
5	XOutDir		1	5	5.4095	.	0.0200
6		TUT	1	4	.	3.6505	0.0561

Deviance and Pearson Goodness-of-Fit Statistics

Criterion	DF	Value	Value/DF	Pr > ChiSq
Deviance	99	103.0607	1.0410	0.3700
Pearson	99	91.7697	0.9270	0.6841

Number of unique profiles: 104

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.1119	0.7555	16.9685	<.0001
RSHF	1	-0.0631	0.0183	11.8752	0.0006
SepChair	1	-1.5927	0.5490	8.4166	0.0037
XOutDir	1	-0.3472	0.1532	5.1361	0.0234
DirShHo	1	-0.0517	0.0199	6.7529	0.0094

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
RSHF	0.939	0.906	0.973
SepChair	0.203	0.069	0.596
XOutDir	0.707	0.523	0.954
DirShHo	0.950	0.913	0.987

Association of Predicted Probabilities and Observed Responses

Percent Concordant	82.5	Somers' D	0.651
Percent Discordant	17.4	Gamma	0.652
Percent Tied	0.1	Tau-a	0.329
Pairs	2704	c	0.826

Profile Likelihood Confidence
Interval for Parameters

Parameter	Estimate	95% Confidence Limits	
Intercept	3.1119	1.7702	4.7631
RSHF	-0.0631	-0.1028	-0.0306
SepChair	-1.5927	-2.7414	-0.5656
XOutDir	-0.3472	-0.6827	-0.0827
DirShHo	-0.0517	-0.0941	-0.0153

Profile Likelihood Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.939	0.902	0.970
SepChair	1.0000	0.203	0.064	0.568
XOutDir	1.0000	0.707	0.505	0.921
DirShHo	1.0000	0.950	0.910	0.985

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
RSHF	1.0000	0.939	0.906	0.973
SepChair	1.0000	0.203	0.069	0.596
XOutDir	1.0000	0.707	0.523	0.954
DirShHo	1.0000	0.950	0.913	0.987

APPENDIX G

THE SOFTWARES USED IN THE STUDY

The simulator used for the implementing the Artificial Neural Network modelling was NEUframe Professional, version 3.0 - 1997 Neural Computer Sciences (NCS).

The Logistic Regression analyses were performed by using SAS, release 8.2 – 1999. SAS Institute Inc.

References

Acquisitions Monthly, (1985-1997), Tudor House Publications Limited, England.

Aggarwal R., Jaffe J. F., Mandelker G.N., (1992),"The Post-Merger Performance of Acquiring Firms: A Re-examination of an Anomaly", *The Journal of Finance*, 47(4), 1605-1621.

Aggarwal A., Navratil F., (1991),"Planning A Successful Acquisition Strategy: Some Guidelines", *Managerial Finance*, 17(1), 14-18.

Ahmadi H., (1990), "Testability of the Arbitrage Pricing Theory by Neural Networks", *Proceedings of the IEEE International Conference on Neural Networks*, 385-393.

Alici, Y., (1996), "Neural Networks in Corporate Failure Prediction: The UK Experience" in Refenes, P., Abu-Mostafa, Y., Moody, J.E., and Weigend, A.S., (Eds), Neural Networks in Financial Engineering (Proceedings of Neural Networks in the Capital Markets (NNCM'95), London, October 1995), World Scientific, Singapore, 1996.

Altman, E., (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*, (September), 23(4), 589-609.

Altman, E., Eisenbeis, (1978), "Financial Applications of Discriminant Analysis: A Clarification"; *Journal of Financial Analysis*, 3, 185-200..

Altman, E., Avery, R., Eisenbeis, R., Sinkey, J.,(1981), "Application of Classification Techniques in Business, Banking and Finance",*JAI*, Greenwich, CT.

Altman, E., Haldeman, R., Narayanan, P., (1977), "Zeta Analysis",*Journal of Banking and Finance*, 1, 29-54.

Altman, E.I., Marco, G., Varetto, F., (1994), "Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks (the Italian Experience)", *Journal of Banking and Finance*, 18(3), 505-529.

Ambrose, B.W., Megginson, W.L., (1992), "The Role of Asset Structure, Ownership Structure, and Takeover Defenses in Determining Acquisition Likelihood", *Journal of Financial and Quantitative Analysis*, 27(4), 575-589.

Amihud, Y., and Lev, B., (1981), "Risk Reduction as a Managerial Motive for Conglomerate Mergers,"*The Bell Journal of Economics*, 12, No. 2, 605-617.

Argenti, J., (1977), "Company Failure-Long-Range", *Accountancy*, August, 46-52.

Asquith P., (1983), "Merger Bids, Uncertainty, And Stockholder Returns",*Journal of Financial Economics*, 11 (1983) 51-83.

Asquith P., Bruner R.F., Mullins D.W., (1983), "The Gains To Bidding Firms from Merger", *Journal of Financial Economics*, 11 (1983) 121-139.

Azoff, E.M., (1994), "Neural Network Time Series Forecasting of Financial Markets", John Wiley & Sons, Chichester.

Baesteans, D.E., Van Den Berg, W.M., Wood, D., (1994), "Neural Network Solutions for Trading in Financial Markets", Financial Times, Pitman Publishing, London.

Barker, D., (1993), "Analysing Financial Health: Integrating Neural Networks and Expert Systems", in Trippi, R. and Turban, E., (Eds), (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Barnes P., (1987), "The Analysis and Use of Financial Ratios: A Review Article", *Journal of Business Finance & Accounting,* 14(4), 449-461.

Barnes P., (1990), "The Prediction Of Takeover Targets In The U.K. By Means Of Multiple Discriminant Analysis", *Journal of Business Finance & Accounting,* 17(1) Spring 1990.

Barnes P., (1998), "Why Do Bidders Do Badly out of Mergers? Some UK Evidence", *Journal of Business Finance and Accounting,* 25(5) & (6), 571-594.

Bartley J.W and Boardman C.M, (1990), "The Relevance of Inflation Adjusted Accounting Data to the Prediction of Corporate Takeovers", *Journal of Business Finance & Accounting,* 17(1) Spring 1990.

Beale R., Jackson T., (1990), *"Neural Computing: An Introduction"* IOP Publishing Ltd, London.

Beaver, W.H., (1966), "Financial Ratios as Predictors of Failure", *Empirical Research in Accounting Selected Studies,* Supplement to *Journal of Accounting Research,* 71-111.

Belkaoui A., (1978), "Financial Ratios As Predictors Of Canadian Takeovers", *Journal of Business Finance & Accounting,* 5(1) 1978.

Bell, T., Ribar, G., Verchio, J., (1990), "Neural Networks vs. Logistic Regression in Predicting Bank Failures", in P.Srivasta (Ed), Auditing Symposium X, University of Kansas.

Bendel, R.B., Afifi, A.A., (1977), "Comparison of Stopping Rules in Forward Regression", *Journal of the American Statistical Association*, 72, 46-53.

Bergerson, K., Wunsch, D.C., (1991), "A commodity Trading Model Based on a Neural_expert System Hybrid", *Proceedings of the IEEE International Conference on Neural Networks,* 289-293.

Berle A. A., Means G. C., (1932), "*The Modern Corporation and Private Property*", Chicago.

Berry, R.H., Triquerios, D., (1993), "Applying Neural Networks to the Extraction of Knowledge from Accounting Reports: A Classification Study", in Trippi, R. and

Turban, E., (Eds), (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Bishop C. M., (1995), "*Neural Networks for pattern Recognition*", Clarendon Press, Oxford.

Blake, D., (1990), "*Financial Market Analysis*", McGraw-Hill, England.

Bosarge, W.E., (1991), "Adaptive Processes to Exploit the Non-Linear Structure of Financial Markets", in Trippi, R. and Turban, E., (Eds), (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Boyd, B.K., (1995), "CEO Duality and Firm Performance: A Contingency Model", *Strategic Management Journal,* 16, 301-321.

Brealey R. A., Myers S.C., (1991), "*Principles of Corporate Finance*", The McGraw-Hill Companies, USA.

Brealey, R.A., Myers, S.C., (1991), "*Principles of Corporate Finance*", McGraw-Hill Inc., USA.

Brickley J. A., Coles J. L., Jarrell G., (1997), "Leadership Structure: Separating the COE and Chairman of the Board", *Journal of Corporate Finance,* 3, 189-220.

Brickley J. A., James C. M., (1987), "The Takeover Market, Corporate Board Composition, And Ownership Structure: The Case Of Banking", *Journal of Law &Economics*, 30, April 1987, 161-181.

Byrd J., Parrino R., and Pritsch G., (1998), "Stockholder-Manager Conflicts and Firm Value", *Financial Analysts Journal,* May-June, 14-30.

Central Statistics Office, (1995), "*Share Ownership*", HMSO, London.

Chiplin, B., Wright, M., (1987), "*The Logic of Mergers*", The Institute of Economic Affairs, London.

Clark, K., Ofek, E., (1994), "Mergers as a Means of Restructuring Distressed Firms: An Empirical Investigation", *Journal of Financial and Quantitative Analysis,* 29(4), 541-565.

Coackley, J.R., Brown, C.E., (1993), "Artificial Neural Networks Applied to Ratio Analysis in the Analytical Review Process", *Intelligent Systems in Accounting, Finance and Management,* 2(1), 19-39.

Collard, J.E., (1993), "Commodity Trading with a Three Year Old", in Trippi, R. and Turban, E., (Eds), (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Collins, E., Ghos, S., Scofield C., (1988), "An Application of a Multiple Neural Network Learning System to Emulation of Mortgage Underwriting

Judgements", *Proceedings of the IEEE International Conference on Neural Networks,* 459-466.

Committee on Corporate Governance (1998), "*Final Report*", January, GEE Publishing, Great Britain.

Conn, R.L., (1985), "A Re-examination of Merger Studies that Use the Capital Asset Pricing Model Methodology", *Cambridge Journal of Economics*, 9, 43-56.

Copeland, T.E., Weston, J.F., (1988), "*Financial Theory and Corporate Policy*", Addison_Wesley Publishing Comp., USA.

Cosh, A.D., Hughes, A., Sing, A., (1980), "The Causes and Effects ofTakeovers in the UK: An Empirical Investigation for the Late 1960s at the Micro-ecenomic Level", In The Determinants and Effects of Mergers (ed. Mueller, D.C) Cambridge.

Cosslett, S.R., (1981), "Maximum Likelihood Estimator for Choice-Based Samples", *Econometrica,* 49, 1289-1316.

Cun Y.L., Denker J.S., Solla S.A., (1990), "Optimal Brain Damage", *Advances in Neural and Information Processing Systems,* 2, 598-605.

Curram S. P., Mingers J., (1994), "Neural Networks, Decision Tree Induction and Discriminant Analysis: An Empirical Comparison", *Journal of Operational Research Society,* 45(4), 440-450.

Cybenko G., (1989), "Approximation by Superpositions of a Sigmoid Function",*Math. Control, Signals & Systems*, 2, 304-314.

Daily C. M., Dalton D. R., (1994), "Bankruptcy and Corporate Governance: The Impact of Board Composition and Structure", *Academy of Management Journal,* December, 37(6), 1603-1618.

Davis, G.F., (1991), "Agents without Principles? The Spread of the Poison Pill through the Intercorporate Network", *Administrative Science Quarterly*, 36, 583-613.

De Carvalho M. C. M., Dougherty M. S., Fowkes A. S., Wardman M. R., (1998), "Forecasting Travel Demand: A Comparison of Logit and Artificial Neural Network Methods", *Journal of the Operational Research Society*, 49, 717-722.

Dickerson A. P., Gibson H. D., and Tsakalotos E., (1998), "Takeover Risk and the Market for Corporate Control: The Experience of British Firms in the 1970s and 1980s", *Department of Economics Discussion Paper 98/3, University of Kent Canterbury*, January.

Dietrich, J.K., Sorensen, E., (1984), "An Application of Logit Analysis to Prediction of Merger Targets", *Journal of Business Research,* 12, 393-402.

Dodd P., (1980), "Merger Proposals, Management Discretion and Stockholder Wealth", *Journal of Financial Economics,* 8, 105-137.

Dodd, P., R. S. Ruback, 1977, "Tender Offers and Stockholder Returns: An Empirical Analysis," *Journal of Financial Economics*, 5: 351-373.

Dutta, S., Shekar S., (1988), "Bond Rating: A Non-Conservative Application of Neural Networks", *Proceedings of the IEEE International Conference on Neural Networks,* 2, 443-450.

Dutta, S., Wong, W.Y., (1994), "Decision Support in Non-Conservative Domains: Generalisation with Neural Networks", *Decision Support Systems,* 11(5), 527-544.

Eisenbeis R. (1977), "Pitfalls In The Application Of Discriminant Analysis In Business, Finance And Economics", *The Journal of Finance,* 32(3) June 1977.

Eisenhardt, K.M., (1989), "Agency Theory: An Assessment and Review",*Academy of Management Review,* 14(1), 57-74.

Ezzamel M., Mar-Malinero C., Beecher A., (1987), "On the Distributional Properties of Financial Ratios", *Journal of Business Finance and Accounting*, 14(4), 463-481.

Ezzamel M., Mar-Malinero, (1990), "The Distributional Properties of Financial Ratios in UK Manufacturing Companies", *Journal of Business Finance and Accounting*, 17(1), 1-29.

Fama E.F., (1980), "Agency Problems and the Theory of the Firm",*Journal of Political Economy,* 88(21), 288-307.

Fama E.F., Jensen M.J., (1983), "Agency Problems And Residual Claims",*Journal of Law &Economics*, 26, June 1983, 327-349.

Fama E.F., Jensen M.J., (1983), "Separation Of Ownership And Control",*Journal of Law &Economics*, 26, June 1983, 301-325.

FAME, Jordans, (1985-1995), Bureau Van Dijk. London.

Feldman K., Kingdon J., (1995), "Neural Networks and some Applications to Finance", *Applied Mathematical Finance,* 2, 17-42.

Firth M., (1980), "Takeovers, Shareholder Returns, and The Theory Of The Firm",*The Quarterly Journal of Economics,* March 1980, 235-260.

Firth M., (1991), "Takeovers, Stockholder Returns, and Executive Rewards", *Managerial and Decision Economics,* 12, 241-248.

Fisher, R.A., (1936), "The Use of Multiple Measures in Taxonomic Problems",*Ann. Eugenics*, 7, 179-188, in Yoon, Y., Swales, G., Margavio, T. M., (1993), "A

Comparison of Discriminant Analysis Versus Artificial Neural Networks", *Journal of Operational Research Society,* 44 (1), 51-60.

Forbes W., Watson R., (1993), "Managerial Remuneration And Corporate Governance: A Review Of The Issues, Evidence And Cadbury Committee Proposals", *Accounting and Business Research,* 23(91A), 1993, 331-338.

Franks J., Mayer C., (1996), "Hostile Takeovers and the Correction of Managerial Failure", *Journal of Financial Economics*, 40, 163-81.

Gammie B., Gammie E., (1996), "The Impact Of Corporate Governance And Performance On Target Selection In A Difficult Macroeconomics Operating Environment", *Proceeding of BAA National Conference,* 27-29th March 1996.

Gately E., (1996), "*Neural Networks for Financial Forecasting*", John Wiley & Sons Inc., USA.

Gerald D.F. and Stout S.K., (1992), "Organization Theory and The Market For Corporate Control: A Dynamic Analysis Of The Characteristics Of Large Takeover Targets, 1980-1990", *Administrative Science Quarterly,* 37 (1992):605-633.

Gort M., (1969), "An Economic Disturbance Theory of Mergers", *The Quarterly Journal of Economics*, February, 73(1), 624-642.

Gregory A., (1997), "The Long-Run Performance of UK Acquirers: Motives Underlying the Method of Payment and Their Influence upon Subsequent Performance", Discussion paper, The Exeter University, Exeter.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., (1995), '*Multivariate Data Analysis with Readings*", Prentice-Hall International Inc., USA.

Hand D. J., (1997), "*Construction and Assessment of Classification Rules*", John Wiley & Sons, England.

Haykin, S., (1999), "*Neural Networks*", Prentice Hall, USA.

Hermalin B.E., Weisbach M.S., (1991), "The Effects Of Board Composition And Direct Incentives On Firm Performance", *Financial Management,* Winter 1991, 101-112.

Hill C. W. L., Jones T. M., (1992), "Stakeholder-Agency theory", *Journal of Management Studies*, 29(2), 127-154.

Holland K.M. and Hodgkinson L., (1994), "The Pre-Announcement Share Price Behavior Of UK Takeover Targets", *Journal of Business Finance & Accounting,* 21(4) June 1994.

Hornik K., Stinchcombe M., and White H., (1989), "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks*, 2, 359-366.

Hosmer D. W., Lemeshow S., (1989), "*Applied Logistic Regression*", John Wiley & Sons Inc., USA.

Hsiao C., (1986), "*Analysis of Panel Data*", Cambridge University Press, Cambridge.

Hsieh S.J., (1993), "A Note on the Optimal Cutoff Point in Bankruptcy Prediction Models", *Journal of Business Finance and Accounting,* 20 (3), 457-464.

Hutchinson, R., (1995), "*Principles of Investment, Financing and Valuation*", Stanley Thornes (Publishers) Ltd., Great Britain.

Jarrell, G., Bradley, M., (1980), "The Economic Effects of Federal and State Regulations of Cash Tender Offers", *Journal of Law and Economics,* October, 371-407.

Jarrell, G., Poulson, A., (1987), "Shark Repellents and Stock Prices: The Effects of Antitakeover Amendments since 1980", *Journal of Financial Economics,* 17, 127-168.

Jensen M.C., (1986), "Agency Costs Of Free Cash Flow, Corporate Finance, And Takeovers", *American Economic Review,* 76, May 1986, 323-329.

Jensen M.C., (1986), "The Takeover Controversy: Analysis and Evidence", *Midland Corporate Finance Journal",* 4, 6-32.

Jensen, M. C. (1993) "The modern Industrial Revolution, Exit, and the Failure of Internal Control Systems", *Journal of Finance* 48, 831-880.

Jensen, M.C., (1994), "Takeovers: Their Causes and Consequences", 15-43, in Gaughan, P.A., (1994), "*Readings in Mergers and Acquisitions*", Basil Balackwell Ltd., USA.

Jensen, M.C., Meckling W.H., (1976), "Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure", *Journal of Financial Economics,* 3, 305-360.

Jensen, M.C., Ruback, R.S., (1983), "The Market for Corporate Control: The Scientific Evidence", *Journal of Financial Economics,* 11, 5-50.

Jensen M.C., Murphy K.J., (1990), "Performance Pay And Top-Management Incentives", *Journal of Political Economy,* 98(2), 1990, 225-265.

Kennedy D. B., (1992), "Classification Techniques in Accounting Research: Empirical Evidence of comparative Performance", *Contemporary Accounting Research,* 8(2), 419-442.

Kennedy D., Lakonishok J., Shaw W.H., (1992), "Accommodating Outliers and Nonlinearity in Decision Models", *Journal of Accounting, Auditing & Finance,* 7, 161-190.

Kennedy V. A., Limmack R. J., (1996), "Takeover Activity, CEO Turnover, and the Market for Corporate Control", *Journal of Business, Finance and Accounting,* 23(2), 267-285.

Kimoto, T., Asakawa, K., Yoda, M, Takeoka, M., (1990), "Stock Market Prediction System with Modular Neural Networks", *Proceedings of the IEEE International Conference on Neural Networks,* 1-16.

Klecka W. R., (1980), "*Discriminant Analysis*", Sage Publication Inc., USA.

Kleinbaum D. G., (1994), '*Logistic Regression: A Self Learning Text* Springer-Verlag New York, Inc. USA.

Kleinbaum D. G., Kupper, L.L., Muller, K.E., Nizam, A., (1998), " Brooks/Cole Publishing Company, USA.

Kryzanowski, L., Galler, M., (1995), "Analysis of Small Business Financial Statements Using Neural Nets", *Journal of Accounting, Auditing and Finance,* 10(1), 147-169.

Langetieg T. C., (1978), " An Application of a Three-Factor Performance Index to Measure Stockholder Gains from Merger", *Journal of Financial Economics,* 6, 365-383.

Lawrence S. R., (1982), "The Application of Analysis of Variance to Inter-Firm Comparison Ratios" *Journal of Business Finance and Accounting,* 9(4), 523-458.

Levine, P., Aaronovitch, S., (1981), "The Financial Characteristics of Firms and Theories of Merger Activity", *Journal of Industrial Economics,* 30, 149-172.

Liang, T., Chandler, J., Han, I., Roan, J., (1992), "An Empirical Investigation of Some Data Effects on the Classification Accuracy of Probit, ID3, and Neural Networks", *Contemporary Accounting Research,* (Fall), 9(1), 306-328.

Lippmann R. P., (1987), "An Introduction to Computing with Neural Nets", *IEEE Acoustics, Speech, and Signal Processing Magazine (ASSP),* April, 4(2), 4-22.

Lumby S., Jones C., (1999), "*Investment Appraisal and Financial Decisions*", International Thomson Business Press, London.

Luther R. K., (1996), "An Artificial Neural Network Approach to Predicting the Outcome of Chapter 11 Bankruptcy", *Journal of Business and Economic Studies*, 4(1), 57-73.

Maddala, G. S. (1983), "*Limited-dependent and Quantitative Variables in Econometrics*" Cambridge University Press, Cambridge.

Maddala, G. S. (1989), '*Introduction to Econometrics*' Macmillan Publishing Company, New York.

Malette P., Fowler K.L., (1992), "Effects of Board Composition and Stock Ownership on the Adoption of Poison Pills", *Academy of Management Journal,* 35, 1010-1035.

Mandelker G., (1974), "Risk and Return: The Case of Merging Firms", *Journal of Financial Economics,* 1, 303-335.

Manne, H. G., (1965) "Mergers and the Market for Corporate Control", *Journal of Political Economy,* February-December, 110-120.

Manski, C.F., Lerman, S.R., (1977), "The Estimation of Choice Probabilities from Choice-Based Samples", *Econometrica,* 45, 1977-1988.

Manski, C.F., McFadden, D., (1981), "*Structural Analysis of Discrete Data with Econometric Applications*", The MIT Press, USA.

Maren A. J., (1990), "Hybrid and Complex Networks", in Handbook of Neural Computing Applications., Maren A. J., Harston C. T. and Pap R. M., (Eds), Academic Press Inc., 1990. pp. 203-219.

Mar-Molinero C., Ezzamel M., (1989), "Initial Data Analysis of Financial Ratios", University of Southampton, Discussion Paper.

Mar-Molinero C., Ezzamel M., (1991), "Multidimensional Scaling Applied to Corporate Failure", *OMEGA International Journal of Management Science,* 19(4), 259-274.

Martin P., (1997), "*Merger Gambit*", Financial Times, (04.09.1997), p24

Martin-del-Brio B., Serrano-Cinca C., (1995), "Self-organizing Neural Networks: The Financial State of Spanish Companies", in Refenes. A.N., (Ed.) (1995), Neural Networks in the Capital Markets, Wiley & Sons, Chichester.

Matsusaka, J. G., (1993), "Takeover Motives during the Conglomerate Merger Wave", *Rand Journal of Economics,* 24, No. 3, 357-379.

McLaney, E.J., (1997), "*Business Finance: Theory and Practice*", Pitman Publishing, Great Britain.

McLeay S., (1986), "Student's t and the Distribution of Financial Ratios", *Journal of Business Finance & Accounting,* 13(2) Summer, 209-222.

Meeks, G., Whittington, G., (1975), "Directors' Pay, Growth and Profitability", *Journal of Industrial Economics,* 24(1), 1-14.

Menard, S., (1995), "*Applied Logistic Regression Analysis, Quantitative Applications in the Social Sciences, Vol 106*", Sage Publications, USA.

Mikkelson W. H., Partch M. M., (1997), "The decline of takeovers and disciplinary managerial turnover", *Journal of Financial Economics*, 44, 205-228.

Miranda F.G., Burgess N., (1997), "Modelling Market Volatilities: The Neural Network Perspective", *The European Journal of Finance*, 3, 137-157.

Mitchell, M. L., Mulherin J. H., (1996), "The Impact of Industry Shocks on Takeover and Restructuring Activity", *Journal of Financial Economics*, 41: 193-229.

Morck, R., Shleifer, A., Vishny, R., (1988), "Management Ownership and Market Valuation", *Journal of Financial Economics*, 20: 293-315

Moschandreas M., (1994), "*Business Economics*", Routledge, USA.

Mozer M.C., Smolensky P., (1989), "Skeletonization: A Technique for Trimming the Fat from a Network via Relevance Assessment", ", *Advances in Neural and Information Processing Systems*, 1, 107-115.

Mueller, D. C., (1969), "A Theory of Conglomerate Mergers", *Quarterly Journal of Economics*, 73(1), February, 643-659.

Mueller, D. C., (1989), "Mergers: Causes, Effects and Policies", *International Journal of Industrial Organization*, 7, 1-10.

NCS, (1997), "Getting Started NeuFrame", Neural Computer Sciences.

Odom, M., Sharda, R., (1990), "A Neural Network Model for Bankruptcy Prediction", *Proceedings of the IEEE International Conference on Neural Networks*, 163-168.

Ohlson, J. A., (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, 18(1), 109-131.

Palepu K.G. (1986), "Predicting Takeover Targets: A Methodological And Empirical Analysis", Journal of Accounting and Economics, 8(1986) 3-35.

Park S., Song M.H., (1995), "Employee Stock Ownership Plans, Firm Performance, And Monitoring By Outside Blockholders", *Financial Management*, 24(4), Winter 1995, 52-65.

Post A. M., (1994), *"Anatomy of a Merger: The Causes and Effects of mergers and Acquisitions"*, Prentice-Hall Inc., USA.

Pound, J., (1988), "Proxy Contests and the Efficiency of Shareholder Oversight", *Journal of Financial Economics*, 20, 237-265.

Powel R. G., (1997), "Modelling Takeover Likelihood", *Journal of Business Finance and Accounting*, 24(7) & (8), 1009-1029.

Price Waterhouse Corporate Register, (1989-1995), Hemmington Scott Financial Publishing Ltd, London.

Rechner P. L., Dalton D. R., (1991), "CEO Duality and Organizational Performance: A Longitudinal Analysis", *Strategic Management Journal,* 12, 155-160.

Refenes, A.N. (Ed), (1995), "Neural Networks in the Capital Markets" ", John Wiley & Sons, Chichester.

Refenes, A.N., (1993), "Constructive Learning and its Applications to Currency Exchange Rate Forcasting", in Trippi, R. and Turban, E., (Eds) (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Refenes, A.N., Zapranis, A., Francis, G., (1994), "Stock Performance Modelling Using Neural Networks: A Comparative Study with Regression Models", *Neural Networks*, 7(2), 375-388.

Rege, U.P., (1984), "Accounting Ratios to Locate Take-over Targets", *Journal of Business Finance and Accounting,* 11(2), 301-311.

Reilly, D.L., Collins, E., Scofield C., Ghos, S., (1991), "Risk Assessment of Mortgage Applications with a Neural Network System: An Update as the Test Portfolio Ages", *Proceedings of the IEEE International Conference on Neural Networks,* 482-497.

Roll, R., (1994), "The Hubris Hypothesis of Takeovers", 44-61, in Gaughan, P.A., (1994), "*Readings in Mergers and Acquisitions*", Basil Balackwell Ltd., USA.

Rosenstein, S., Wyatt, J. G., (1997), "Inside directors, board effectiveness, and shareholder wealth", *Journal of Financial Economics,* 44, 229-250.

Ruback R., (1983), "Assessing Competition In The Market For Corporate Acquisitions", *Journal of Financial Economics,* 11 (1983) 141-153.

Rumelhart D.E., McClelland J.L. and the PDP Research Group (1992), "*Parallel Distributed Processing: Explorations in the Microstructure of Cognition*", Volume I: Foundations, MIT Press, London.

SAS Enterprise Miner for Windows NT, (1999), "Neural Network Node", SAS Institute Inc., USA.

SAS System, (1995), "Logistic Regression Examples", SAS Institute Inc., USA.

Schalkoff R., (1992), "*Pattern Recognition: Statistical, Structural and Neural Approaches*", John Wiley & Sons, Inc., USA.

Scherer, F. M., (1988), "Corporate Takeovers: The efficiency arguments", *Journal of Economic Perspectives,* 2(1), Winter, 69-82.

Sen T.K., Oliver R., and Sen N., (1995), "Predicting Corporate Mergers", *Neural Networks in the Capital Markets,* John Wiley & Sons Ltd.

Serrano-Cinca C., (1997), "Feedforward Neural Networks in the Classification of Financial Information" *The European Journal of Finance*, 3, 183-202

Sharda, R., Patil, R.B., (1993), "A Connectionist Approach to Time Series Prediction-An Empirical Test", in Trippi, R. and Turban, E., (Eds) (1993), Neural Networks in Finance and Investing, Probus Publishing, London.

Shivdasani A., (1993), "Board composition, ownership structure, and hostile takeovers", *Journal of Accounting and Economics,* 16, 167-198.

Shleifer A., Vishny R.W., (1986), Large Shareholders And Corporate Control,*Journal of Political Economy,* 94(3), 1986, 461-488.

Simkowitz, M., Monroe, R.J., (1971), "A Discriminant Analysis Function for Conglomerate Targets", *Southern Journal of Business,* November, 1-16.

Singh, A., (1971), "*Take-overs, Their Relevance to the Stock Market and the Theory of the Firm*" Cambridge University Press, Cambridge.

Singh, A., (1975), "Take-overs, Economic Natural Selection, and the Theory of the Firm: Evidence From The Postwar United Kingdom Experience", *The Economic Journal,* 85, 497-515.

Singleton, J.C., (1990), "Neural Nets for bond rating improved by multiple hidden layers", *Proceedings of the IEEE International Conference on Neural Networks,* 151-162.

Smith, A., (1776), "*An Inquiry into the Nature and Causes of the Wealth of Nations*" in Moschandreas M., (1994), "*Business Economics*", Routledge, USA.

Smith, T., (1992*), "Accounting for Growth: Stripping the Camouflage from Company Accounts*", Century Business, London.

So J. C., (1987), "Some Empirical Evidence on the Outliers and the Non-normal Distribution of Financial Ratios", *Journal of Business Finance & Accounting,* 14(4) Winter, 483-496.

Song, M.H., Walking, R.A., (1993), The Impact of Managerial Ownership on Acquisition Attempts and Target Shareholder Wealth",*Journal of Financial and Quantitative Analysis,* 28(4), 439-457.

Stapledon, G.P., (1996), "*Institutional Shareholders and Corporate Governance*", Clarendon Press, Oxford.

Steiner M., Wittkemper H., (1997), "Portfolio Optimization with a Neural Network Implementation of the Coherent Market Hypothesis", *European Journal of Operational Research*, 100, 27-40.

Stevens, D. L., (1973), "Financial Characteristics of Merged Firms: A Multivariate Analysis", *Journal of Financial and Quantitative Analysis*, 8, 149-165.

Stultz, R.M., Johnson, H., (1985), "An Analysis of Secured Debt",*Journal of Financial Economics*, 14, 501-521.

Stultz, R.M., (1988), "Managerial Control of Voting Rights: Financing Policies and the Market for Corporate Control", *Journal of Financial Economics*, 20, 25-54.

Sudarsanam P. S., (1985), "Industrial Classification in UK Capital Markets: A Test of Economic Homogeneity", *Applied Economics*, 17, 291-308.

Sudarsanam P. S., (1995), "*The Essence of Mergers and Acquisitions*", Prentice-Hall International, London.

Sudarsanam S., Holl P., Salami A., (1996), "Shareholder Wealth Gains in Mergers: Effect of Synergy and Ownership Structure", *Journal of Business Finance and Accounting*, 23(5), 673-698.

Sudarsanam S., (1991), "Defensive Strategies of Target Firms in U.K. Contested Takeovers", *Managerial Finance*, 17(6), 47-56.

Sudarsanam S., Taffler R.J., (1995), "Financial Ratio Proportionality and Inter-temporal Stability: An Empirical Analysis",*Journal of Banking and Finance*, 19, 45-60.

Taffler R.J., (1982), "Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data", *Journal of Royal Statistical Association"*, 145, Part 3, 342-358.

Tam, K. Y., Kiang, M. Y., (1992), "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions",*Management Science*, 38(7), 926-947.

Tam, K.Y., (1991), "Neural Network Models and the Prediction of Bank Bankruptcy", *OMEGA International Journal of Management Science*, 19(5), 429-445.

Tarrasenko, L., (1998), "*A Guide to Neural Computing Applications*", John-Wiley & Sons Inc., New York.

The Bank of England, (1983), "The Composition Of Company Boards In 1982", *Bank of England Quarterly Bulletin*, 23(1), March, 66-68.

The Bank of England, (1985), "The Boards Of Quoted Companies", *Bank of England Quarterly Bulletin*, *25*(2), June, 233-236.

The Cadbury Report, (1992).

Tzoannos, J., Samuels, J.M., (1972), "Mergers and Takeovers: The Financial Characteristics of Companies Involved", *Journal of Business Finance*, 1, 5-16.

Udo, G., (1993), "Neural Network Performance on the Bankruptcy Classification Problem", *Computers and Industrial Engineering*, 25, 377-380.

Vaughn M. L., (1996) "Interpretation and Knowledge Discovery from the Multilayer Perceptron Networks: Opening the Black Box", *Neural Computing & Applications*, 4(2), 72-82.

Walkling R.A, (1985), "Predicting Tender Offer Success: A Logistic Analysis", *Journal of Financial and Quantitative Analysis*, 20(4) December 1985.

Walkling, R.A., Long, M.S., (1984), "Agency Theory, Managerial Welfare, and Takeover Bid Resistance", *The Rand Journal of Economics*, 15(1), 54-68.

Walter, R.M., (1994), "The Usefulness of Current Cost Information for Identifying Takeover Targets and Earning Above-Average Stock Returns", *Journal of Accounting Audit. Finance*, 9, 349-377.

Wansley, J.W., Lane W.L., (1983), "A Financial Profile of Merged Firms", *Review of Business & Economic Research*, Fall, 87-98.

Wasserman P. D., (1993), "*Advanced Methods in Neural Computing*", Van Nostrand Reinhold, New York.

Watson C.J., (1990), "Multivariate Distributional Properties, Outliers, And Transformation of Financial Ratios, *The Accounting Review*, 65(3), July 1990, 682-695.

Weir C., (1997), "Corporate Governance, Performance and Takeovers: An Empirical Analysis of UK Mergers". *Applied Economics*, 29, 1465-1475.

Weisbach, M.S., (1988), "Outside Directors and CEO Turnover", *Journal of Financial Economics*, 20 (1988) 431-460.

Weiss, S.M., Kulikowski, C.A., (1991), "*Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems*", Morgan Kaufmann Publishers Inc, San Mateo.

Whittington G., (1993), "Corporate Governance and the Regulation of Financial Reporting", *Accounting and Business Research*, 23(91A), 1993, 311-319

Williamson, O.E., (1964), "*The Economics of Discretionary Behaviour: Managerial Objectives in a Theorf of the Firm*", Prentice Hall.

Wilson, R.L., Sharda, R., (1994), "Bankruptcy Prediction using Neural Networks", *Decision Support Systems*, 11(5), 445-557.

Yoon, Y., Swales, G., (1991), "Predicting Stock Price Performance: A Neural Network Approach", *Proceedings of the IEEE 24th Annual Hawaii International Conference of System Sciences*, 156-162.

Yoon, Y., Swales, G., Margavio, T. M., (1993), "A Comparison of Discriminant Analysis Versus Artificial Neural Networks", *Journal of Operational Research Society,* 44 (1), 51-60.

Young K., (1999), "A Learning Process", *The Banker,* January, 149, 875, 65-67.

Zavgren C. V., (1983), "The Prediction of Corporate Failure: The State of the Art", *Journal of Accounting Literature,* 2, 1-35.

Zmijewski M. E., (1984), "Methodological Issues Related to the Estimation of Financial Distress Prediction Models", *Journal of Accounting Research,* 22, Supplement 1984, 59-82.