Comparing the predictive performance of the logit and discriminant analysis models in bankruptcy prediction

A thesis submitted in partial fulfilment of the requirements

for the Degree

of Master of Commerce in Accounting

at the University of Canterbury

by Mari Ono

2018

University of Canterbury

Department of Accounting and Information Systems

Senior Supervisor: Dr. Michael Falta

Co-Supervisor: Dr. Warwick Anderson

Abstract

This research assesses the sensitivity of ratio analysis in bankruptcy prediction. I challenge the usually applied criterion in bankruptcy analysis where a model outperforms another model based on their classification performance (prediction) of firms into bankrupt and non-bankrupt categories. The commonly applied decision criterion is based on Type 1 and Type 2 errors or the aggregate predictive ability of a model. However, I claim that, say, if these metrics show identical results for two models, it is not clear by any stretch that they have the same predictive ability. Rather I contest that looking at the classification patterns of each individual firm, one may observe that indeed both models classify rightly and wrongly all firms in a testing sample; or, they do not, in which case a data-driven result has been obtained.

I have chosen to demonstrate this novel idea by means of the classification performance of the multivariate discriminant analysis (MDA) and Logit models. These two models were introduced by Altman (1968) and Ohlson (1980), respectively, and are probably the most established bankruptcy prediction tools in the literature and in business applications. In doing so, I heavily lean on Begley et al. (1996) who compare the two models using data from the 1990ies. From this starting point, I further apply criteria that allow for a fairer comparison of the two models, because testing two models about the predictive ability to classify firms into groups seems unreliable when each model uses different variables and where for neither of the two models any specification test results are considered.

My research question therefore asks whether the reported results of the MDA and Logit models in the literature with respect to bankruptcy prediction do hold when firm-specific classification patterns are considered. I provide various negative results to this question. This work and the discussed future direction that may be taken based on my work have to potential to redirect the bankruptcy literature to more useful assessments of model performance comparisons.

Acknowledgments

Firstly I would like to thank everyone who supported me to complete the Master study. I especially would like to thank my supervisor, Dr Falta, for his support, guidance and encouragement in carrying out the research. I also would like to thank my co-supervisor, Dr Warwick, as well as all my lectures at UC who provided professional knowledge to add to this research.

Secondly, I would like to thank my friends who shared a student time at UC, who supported me in Christchurch, Auckland, Japan, and the US. Without your support, I could not see myself finishing the research.

Last but not the least, I thank my family who always believes in me, inspires me and supports my challenges.

Mari Ono, February 2018

Contents

Chapter 1. Introduction

1.1 Importance of bankruptcy prediction models

Bankruptcy analysis is a classification exercise which predicts whether companies will fail or not. The bankruptcy prediction models which predict such outcomes are used to ascertain credit ratings and probabilities of corporate failure. Are used in both academia and practice. In practice, they are used by financial institutions and by professional investors as the investment decision making tools. Some bankruptcy prediction models are not complicated to use, others require advanced knowledge in statistics and accounting. Also, many websites offer calculation tables which automatically assess the corporate failure probability with limited financial information. As a result, stakeholders who are interested in the bankruptcy prediction models are spread across the world. Any person or institute who makes critical economic decisions on their investment can be a user of the prediction models.

An important purpose for using a bankruptcy model is to assess the credit worthiness of a company. For example, financial institutions have a specific interest in bankruptcy prediction. As their ‘service’ is to supply capital to the companies, their main interest is to assess credit worthiness of the borrowers so that they can control the default risks of loans and maximize profit of their portfolio (Altman, Iwanicz, Drozdowska, Laitinen, & Suvas, 2016). Investors have a similar interest in terms of their portfolio management. They are interested in determining the risk of returns of their investment and minimizing the level of default risks. Another example of using bankruptcy models is in relation to M&A (merger and acquisitions) opportunities. One of the solutions to exit from near-to-bankruptcy situations for distressed companies is M&A but it is up to the acquirer to make the offer. For such companies, bankruptcy prediction models are useful tools to look for diversification, growth and invest options.

Statistical bankruptcy analysis began in the 1960ies due to the development of statistical techniques and computer technology. For example, Beaver (1966, 1968) used univariate analysis for selected ratios and detected that some of these had a useful predictive indicator. Altman (1968) moved the model forward by developing a multiple discriminant analysis model (MDA) called the Z-Score model. Ohlson (1980) used the logit model and Zmijewski (1984) summarized these models and developed his own model taking a probit approach.

More recent bankruptcy analysis studies attempted to adopt financial models and classification techniques. Hillegeist et al. (2004) assessed the bankruptcy risk information by using an option-pricing model. Neophytou & Molinero (2004) employed multidimensional scaling techniques and created a map to visualise the areas where bankrupt and non-bankrupt companies. Other studies use classic bankruptcy models, such as Altman’s Z-score (1980) and the Logit model by Ohlson (1980), and apply them to unique datasets. Baldwin and Glezen (1992) used the quarterly financial statement information and found it more useful than annual financial information to predict bankruptcies. Aly, Barlow & Jones (1992) showed that combining historical cost information and current cost information predicts bankruptcy more accurately than if only the historical cost information for three years before bankruptcy was included.

Progress in the bankruptcy prediction literature is deemed to have ‘occurred’ when the classification errors are compared and a model comes out on top of an existing model. When a study evaluates the classification errors - Type 1 and Type 2 error rates - the research concludes that the model which makes the lower (combined) error rates is better than that the model with higher (combined) error rates. There are many, many such published studies (e.g. Grice & Ingram, 2001; Wu, Gaunt & Gray, 2010). The problem with such a model evaluation approach is that the results show only the aggregate error rate. If, for example, two models show the same error rate (the aggregate level), there still may be differences in how individual companies have been classified. The results obtained are therefore sample dependent which results in tenuous statements about which model has come on top in terms of prediction performance. In other words, if two models have the same error rate but have misclassified different companies, it cannot be concluded which of the two models is superior in detecting bankruptcy. What that means is that the models are more sensitive to different failure modes. Therefore, comparing Type 1 and Type 2 error rates only is not sufficient to conclude the superiority of one model over another. Though there are some studies which disclose misclassified companies (Altman, 1968, Wu, Liang, & Yang, 2008), no study has attempted to critically reveal the evaluation system of classification performance at the firm level. Here, I am going to focus on the problem of the evaluation system of bankruptcy studies and shed light on the validity of the claim that one model is better than another model.

In Chapter 2, I will review some of the important literature on MDA and Logit models. I also critically discuss the issues of bankruptcy prediction relating to the application of these models. In Chapter 3, I explain the problems with the evaluation system of bankruptcy prediction studies, and then I propose my research question in Chapter 4. I explain my methodology and dataset which I used in this research in Chapter 5, and present my analysis in Chapter 6 with a replication of the Begley et al. study and in Chapter 7 where I address the research question in a set of experiments that analyse the predictive ability of the MDA and Logit models. Chapter 8 concludes the research.

Chapter 2. Literature review: bankruptcy prediction models

The literature review consists of two chapters. Chapter 2 reviews some of the empirical studies of applications using multiple discriminant analysis (MDA) and the Logit model. The issues about bankruptcy prediction analysis are discussed from three different aspects; problems with dataset selections, problems with statistical assumptions, and necessities to consider a time dimension. Chapter 3 deals with problems with the evaluation system of bankruptcy prediction. In particular, presenting examples of Type 1 and Type 2 error rates, I am going to explain the shortcomings of the current system to evaluate the prediction ability of models.

2.1 Statistic prediction models: MDA and Logit

The MDA and Logit models have been widely employed in both practice and academic studies (Altman, Iwanicz-Drozdowska, Laitinen & Suvas, 2014). Some studies compare the prediction ability between the two models. Others compare the prediction ability between either the MDA or Logit with other models. In each study, the researcher(s) use unique data samples which lead the research to conclude the prediction ability of the models under study.

2.1.1 MDA

MDA is an extension of a linear discriminant analysis (LDA), which is a statistical technique developed by R.A. Fisher in the 1930ies (Altman, 1968). LDA is used for predicting qualitative values, amongst other things. One of the early applications of MDA in bankrupt study is due to Altman (1968). He observed 66 companies, of which 33 companies are bankrupted. The industry was the manufacturing sector and the time period of the research was from 1946 to 1965. Altman (1968) listed 22 potentially important financial ratios for analysis, and classified these variables into five standard categories: liquidity, profitability, leverage, solvency, and activity ratios. From these categories, he selected the five ratios that best predicted the corporate failure.

The discriminate function is called Z-score

$$\boldsymbol{Z = 0.012X_1 + \ 0.014X_2 + \ 0.033X_3 + \ 0.006X_4 + \ 0.999X_5}, \qquad \text{(eqn. 1)}$$

where

X_1 = Working capital / Total assets,

X_2 = Retained earnings / Total assets,

X_3 = Earnings before interest & taxes / Total assets,

X_4 = Market value equity / book value of total debt,

X_5 = Sales / total assets.

Altman finds that if the Z-score is greater than 2.99, the company is classified into the non-bankrupt group; if the Z-score is less than 1.81, the company is classified into the bankrupt group; and if the Z-score is between 1.81 and 2.99, the company is in the grey area and it is uncertain if the company is going to bankrupt or not. The original Z-score showed higher predictive classification accuracy on predicting bankruptcy: 95 % of all firms in the bankrupt and non-bankrupt groups were classified to their actual group. Type 1 and Type 2 error rates were 6% and 3%, respectively (Table 2-1).

Table 2-1 Type 1 and Type 2 error rate of Altman's study (1968)

Altman (1968)	Number Correct	Percent Correct	Percent Error	Total Number
Type 1 errors	32	97.0%	3.0%	33
Type 2 errors	31	93.9%	6.1%	33
Total error rate		95.5%	4.5%	

Altman (1983) revised his models several times to fit different types of samples. For exmaple, he revised the Z-socre for predicting bankrupcy of private firms (Altman, & Hotchkiss, 2010). Because private firms do not have a market value, he revised the model replacing” market value equity / book value of total debt (X_4)" with "book value of equity/book value of total debt". The revised model is

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.998X_5\,, \qquad \text{(eqn. 2)}$$

where the cutoff score between the non-bankrupt firms and the grey airea is 2.67, the Type 1 error rates for the Z and Z' models are same (3.0%), but the Type 2 error rates increased to 9.1% when using the revised Z' model.

2.1.2 The Logit model

Ohlson (1980) used a logit model to predict corporate bankruptcy. The logit model is based on the nonlinear maximum likelihood method. This approach overcomes some statistical problems of the MDA approach, such as the assumption of Normally distributed error terms and that the a priori relationship between predictors and outcome is linear.

Ohlson used 105 bankrupted companies and 2,058 non-bankrupted companies in the period from 1970 to 1976. Financial ratios employed in his study are as follows:

1. SIZE – In (total assets/GNP price level index). The index assumes a base value of 100 for 1968.
2. TLTA – Total liabilities/Total Assets
3. WCTA – Working capital/ Total assets
4. OLCA – Current liabilities/ Current assets
5. OENEG – One if total liabilities exceed total assets, otherwise zero
6. NITA – Net income/Total assets
7. FUTL – Funds provided by operations/Total liabilities
8. INTWO – One if net income was negative for the last two years, otherwise zero
9. NHIN – $(NI_t - NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$, where NI_t is net income for the most recent period.

Ohlson (1980) compares the prediction ability of the logit model using different forecasting horizons of 1 year and 2 years. Table 1 shows the Type 1 and Type 2 errors at selected cut-off values (P) for 1 year and 2 years forecasting horizons. A higher P value means a higher probability of bankruptcy. Ohlson finds that when P is at 0.038, the sum of Type 1 and Type 2 errors is minimized for the 1 year forecasting horizon (Type 1 error is 17.4% and Type 2 error is 12.4%). When comparing the average error rates of the two horizons, 1 year forecasting horizon performs better at several points. However, when observing Type 1 and Type 2 error rates separately, in overall Type 1 error rates of 1 year forecasting horizon are lower than that of 2 years forecasting horizon, and Type 2 error rates of 2 years forecasting horizon are lower than that of 1 year forecasting horizon. Based on the comparison of error rates, Ohlson concludes that the logit model can be used for predicting bankruptcy in both forecasting horizon. Ohlson also look at a t-statistics of variables and identified four statistically significant factors which influence the probabilities of failure: size of the company, measures of the companies' financial structure, measures of performance, and measures of current liquidity.

Table 2-2 - Ohlson's model

	1 year forecasting horizon			2 years forecasting horizon		
Estimated probability of bankruptcy: Cutoff point (P)	Type 1 error (%)	Type 2 error (%)	Average	Type 1 error (%)	Type 2 error (%)	Average
0.00	100.00	0.00	50.00	100.00	0.00	50.00
0.02	**28.70**	7.60	**18.15**	54.30	**0.00**	27.15
0.04	**16.70**	14.30	**15.50**	37.70	**0.95**	19.33
0.06	**11.80**	20.00	15.90	26.80	**4.76**	**15.78**
0.08	**9.30**	25.70	17.50	20.20	**8.60**	**14.40**
0.10	**7.20**	26.70	16.95	17.00	**12.40**	**14.70**
0.20	**3.30**	44.80	24.05	7.20	**31.40**	**19.30**
0.30	**1.75**	48.60	25.18	3.60	**43.80**	**23.70**
0.40	**1.07**	57.10	29.09	2.00	**50.50**	**26.25**
0.42	**0.92**	61.00	30.96	1.75	**51.40**	**26.58**
0.50	**0.63**	67.60	34.12	1.07	**57.10**	**29.09**
0.54	**0.44**	68.60	34.52	0.82	**61.00**	**30.91**
0.60	**0.29**	71.40	35.85	0.68	**62.90**	**31.79**
0.70	**0.19**	76.20	38.20	0.49	**70.50**	**35.50**
0.80	**0.15**	81.90	41.03	0.24	**74.30**	**37.27**
0.90	**0.05**	88.60	44.32	0.19	**82.90**	**41.55**
1.00	0.00	1.00	0.50	0.00	1.00	0.50

2.1.3 Related studies

Though many different types of bankruptcy prediction models have been introduced since 1968, the MDA and Logit models have been the dominant models to predict a bankruptcy in academic resarch. They have also been employed all over the world in practice (Altman et al., 2014). The researchers have examined the prediction ability of the two models using different type of samples. Sample diffences are usually due to selecting different time periods, countries, and industries, all of which influence to a predictive ability of a model.

1) Time period

The time period is one of the important factors when selecting a sample. Even though the dataset may contain the same financial ratios, if the data are from different time periods, a comparison of the prediciting ability of two models may yield reverse results. For example, the economic situations that may have changed at both macro and micro-levels or based on legal, tax and accounting regulations that may have been amended could drive the result obtained.

Begley, Ming & Watts (1996) re-examine the classificaion errors of the original Z-score and Ohlson's logit models using datasets from the 1980ies. As the datasets of the original studies were selected from the 1940ies through to the 1970ies, there is at least a 10 year gap between the original studies and Begley's study. Begley et al. pointed out that due to the decrease of corporate debt level in the 1980ies and the changes of the bankruptcy law in the 1970ies, they expect the classification errors to differ from the Altmand and Ohlson studies. Begley et al.'s sample was selected from listed companies on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Securities Dealers Automated Quotations (NASDAQ). The time preiod of the study was from 1980 to 1989, and the industry classes were similar to the original studies of the 1960ies (SIC code less than 4000 or between 5000 and 5999).

Table 2-3 shows how the error rates change when testing the original Z-score using a Begley's hold out sample from the 1980ies. It shows that Type 1 and Type 2 errors increased to

18.5% and 25.1%, respectively, and that the total error rates increased from 4.5% to 21.8%, respectively. The results indicate that the prediction accuracy of the original Z-score decreased when a different hold out sample is used.

Table 2-3 Type 1 and Type 2 error rates from applying the original Z-score to 1980s hold out sample vs Altman's result (1968)

Altman (1968)	Number Correct	Percent Correct	Percent Error	Total Number
Type 1 errors	32	97.0%	3.0%	33
Type 2 errors	31	93.9%	6.1%	33
Total error rate		95.5%	4.5%	
Begley et al. (1996)	Number Correct	Percent Correct	Percent Error	Total Number
Type 1 errors	974	74.9%	25.1%	1300
Type 2 errors	53	81.5%	18.5%	65
Total error rate		78.2%	21.8%	

The results of applying Ohlson's model

Table 2-4 shows how the results change when testing the Ohlson model by a Begley et al.'s hold out sample from the 1980ies. At the original cutoff point of 0.038, the Type 1 error rate increased from 17.4% to 26.6%, while the Type 2 error rate slightly decreased from 12.4% to 10.8%. The total error rate increased from 14.9% to 18.7%.[1] Although the Type 2 error rate in Begley's study is slightly lower than Ohlson's, the rate of Type 1 errors and the total error rate of Begley's study noticeably increased. This indicates that the prediction accuracy of the Ohlson's model changes (decrease, in this case) when it is tested by a hold out sample from a different time period.

[1] The total error rate an average of Type 1 and Type 2 errors. For example, Ohlson's total error rate is calculated by the following formula. (17.4%+12.4%)/2 = 12.4%

Table 2-4 - Type 1 and Type 2 error rates of Ohlson's (1980) vs Begley et al.'s (1996)

	Ohlsons's (1980)		Begley et al. (1996)	
Estimated probability of bankruptcy: Cutoff point	Type 1 error (%)	Type 2 error (%)	Type 1 error (%)	Type 2 error (%)
0.00	100.0	0.0	100.0	0.0
0.02	28.7	7.6	38.0	9.2
0.038	17.4	12.4	26.6	10.8
0.04	16.7	14.3	25.5	10.8
0.06	11.8	20.0	19.1	20.0
0.08	9.3	25.7	15.7	26.1
0.10	7.2	26.7	13.1	30.8
0.20	3.3	44.8	7.5	53.8
0.30	1.8	48.6	5.1	58.5
0.40	1.1	57.1	3.5	66.1
0.42	0.9	61.0	na	na
0.50	0.6	67.6	2.5	70.8
0.54	0.4	68.6	na	na
0.60	0.3	71.4	1.8	73.8
0.70	0.2	76.2	1.5	80.0
0.80	0.2	81.9	0.9	90.8
0.90	0.0	88.6	0.5	95.4
1.00	0.0	1.0	0.0	0.0
Cutoff point that minimize Type 1 and Type 2 errors	0.038		0.041	

Prediction accuracy of the Altman and Ohlson models

In summary, the Type 1[2] error rates of Altman and Ohlson models are 25.1% and 26.6%, respectively. The type 2 error rates of Altman and Ohlson models are 18.5% and 10.8%, respectively. As the cost of Type 2 error is expected to be higher than the cost of Type 1 errors Begley et al. concluded that Ohlson model outperforms the Altman model as the total error rates of the Ohlson model were lower than those for the Z-score (Altman model).

2) Country

Corporate bankruptcy has been studied all over the world, and some researchers selected their

[2] Type 1 and Type 2 errors are explained in Chapter 3.

samples from their own country. Such samples reflect a country's unique aspects, such as religion, culture, law and regulations, and the size of the economy.

Back, Laitinen, Sere, & van Wezel (1996) used datasets from randomly selected 37 Finnish bankrupt companies and matched non-bankrupt companies. The time period of the research was between 1986 and 1989. Most of the companies were from the manufacturing industry. As a number of large companies which bankrupted in Finland was small, the selected companies were mainly SMEs.

Back et al. (1996) compared the Type 1 and Type 2 error rates based on applying discriminant analysis, logit analysis, and genetic algorithms (neural networks). Table 2-5 shows that the Type 1 and Type 2 error rates for one year prior to bankruptcy. When comparing the total error rates, the neural networks produced significant lower error rates at 2.7%, while discriminant analysis and logit analysis produced relatively higher error rates at 14.86% and 13.51%, respectively. The research concluded that neural networks outperformed discriminant analysis and logit analysis.

Table 2-5 - Type 1 and Type 2 error rates for DA, Logit analysis and NN

DA			Logit			NN		
Type 1 error	Type 2 error	Total error	Type 1 error	Type 2 error	Total error	Type 1 error	Type 2 error	Total error
13.51%	16.22%	14.86%	13.51%	13.51%	13.51%	5.26%	0.00%	2.70%

(DA: Discriminant analysis, NN: Neural networks)

3) Industry

Industry categorizes firms according to their economic function, their products and services, and their target markets. If industry is different, business models are different, and so would the financial representations of those firms be. Hence variables selected to contain discriminatory power between failed on non-failed companies, and parametric estimation are likely to vary from industry to industry. For example, every industry has characteristic ratio bands.

For example, Ho, McCarthy, Yang, & Ye (2013) focused on the North American pulp and paper industry. Although the industry is one of the most important sectors in the U.S economy, the companies had faced serious declines in demand that resulted in increasing the number of bankrupt firms. Ho et al. re-estimated the Ohlson Logit model (1980) using a datasets of the period between 1990 and 2005, and compared the predictive ability between Ohlson's original model and their re-estimated model by testing an unique testing sample. The testing sample is created from North American publicity listed pulp and paper companies, and consists of 2 bankrupt companies and 42 non-bankrupt companies in 2007 and 2008. Ho et al. also re-estimated Altman's (1968) model and compared the predictive ability between Altman's original model and their re-estimated model by a testing sample dataset

The results of applying Ohlson's model

Table 2-6 shows the comparison between the Type 1 and Type 2 errors for Ohlson's original model and the re-estimated model using the same testing sample dataset. The optimal cutoff point of the original model was 0.038 while the optimal cutoff of the re-estimated model was 0.14. When applying the two models to the testing sample, Ohlson's original model predicted the two bankrupted companies correctly (the Type 1 error rate is 0%). However, the Type 2 error rate is 100%, which indicated that the cutoff point does not minimize the total error rates. In contrast, the re-estimated model did not predict the two bankrupt companies to go bankrupt (the Type 1 error rate is 100%), but the Type 2 error rate of re-estimated model was smaller at 2% compared the Ohlson model. Overall, the total error rates of Ohlson's original model and the re-estimated model were 91% and 7%, respectively. Based on the results, the research concludes that the re-estimated Ohlson model has better prediction ability than the original model.

Table 2-6 - Type 1 and Type 2 errors of Ohlson's and re-estimated model

	Ohlson's model	Re-estimated model
Type 1 error	0.0%	100.0%
Type 2 error	100.0%	2.0%
Total error	91.0%	7.0%

The results of applying Altman's model

Table 2-7 shows the comparison between the Type 1 and Type 2 error rates of Altman's original model and Ho's re-estimated model. Both Altman's and the re-estimated model predicted the two bankrupt companies correctly (both Type 1 error rates are 0%). However, the Type 2 error rate of the original Z-score was 31.0%, which is 24.0% higher for the re-estimated model. The total error rates of the original Altman and re-estimated model were 3.0% and 7.0%, respectively. Based on these results, the researchers concluded that the re-estimated model has better predictive ability.

Table 2-7 - Type 1 and Type 2 error rates of Altman's and re-estimated model

	Altman's model	Re-estimated model
Type 1 error	0.0%	0.0%
Type 2 error	31.0%	7.0%
Total error	3.0%	7.0%

2.2 Problems with bankruptcy analysis

One reason that bankruptcy studies have not been able to find a "best" bankruptcy model is due to the discussed differences in samples. The number of sample selection criteria available for bankruptcy studies is large and left as a matter of choice to the researcher(s). If the samples are different (but all other conditions are same), it is not surprising that the outcomes are different. But which method is then the practitioner to choose? Each study decides about some specific assumptions, such as samples and datasets, adopted models and statistical assumptions. This discretion produces inconsistent classification errors and obscures the creditability of the bankruptcy literature.

2.2.1 Problem with the 'dataset'

There is no universal definition of failure in bankruptcy analysis studies. This is because bankruptcy in practice is closely related to legal systems, regulations, culture and business customs, and they vary from country to country. Some studies use a juridical definition of failure, mostly bankruptcy. Others refer to financial distress, but there are many different ways to describe financial distress, such as bankruptcy, insolvency, loan default, moratorium, liquidation, and government support. For example, Altman (1968) defined that the bankrupt firms are legally bankrupt and either placed in receivership or reorganization. If researcher(s) defined the bankrupt firms based on a different category, the study ought not to compare the accuracy of bankruptcy models with Altman's.

Other ambiguous factors in creating samples in bankruptcy studies include industries, countries and regions, the size and age of organisations, and time and period of research. For example, Altman's (1968) originally focused on the manufacturing industry. More recent studies focus on a financial industry particularly in the wake of financial crises (e.g., Douglas, Lont & Scott, 2014; Iturriaga & Sanz, 2015). Since business models between manufacturing and financial industry companies are different and so are the, say, regulatory requirements of these industries, their financial and non-finanical information will have to comply with (or is subject to) differing constraints. How then can a model claim to be superior over another model if such judgement depends on industry? To put it differently, though the purpose of bankruptcy studies is to find a model which is better than others, conclusions about such models are only valid to the extent of defined samples. As different samples deliver different results, the predictability of the bankruptcy models should not be compared across the industries. Likewise, if datasets are created from cross-industries' information, more sample-driven results will be obtained.

Procedures around the treatment of outliers is a further source of what eventually will result in data-driven conclusions. When outliers are in datasets, researcher tend to trim (Winsorize) them in order to closer conform to the linearity relationship between the dependent variables and the explanatory variables. However, outliers may contain useful information and are paramount to model and theory testing.

2.2.2 Problem with statistical assumptions

Making a linearity assumption is often seen in bankruptcy prediction studies. As it is easier to estimate parameters in a linear model, the simplified model-based approach is often preferred. However, the potential problem of making a linearity assumption is that estimated models do not representationally faithfully describe observable, real data. The real data may show a non-linearity relationship, and in this case it is only in approximation that the models are useful. Yet, the degree of approximation ought to be demonstrated.

Multicollinearity is another problem. Collinearity occurs when two or more predictor variables are closely correlated to one another. Multicollinearity introduces a bias into the parameter estimates which will affect the classification performance of a model. For example, the original Z-score employs five financial ratios and four of them use total assets in the denominator. However, some studies assume that multicollinearity is irrelevant in MDA models (Eisenbeis, 1977; Altman and Eisenbeis, 1978; Balcaen & Ooghe, 2006).

2.2.3 Problem with time dimensions

The majority of statistical bankruptcy prediction studies focus on one annual account and take one single observation in selecting financial ratios from each firm. Bankruptcy models of such studies do not take in account the trend and failure process when predicting failure. This causes a problem called “snapshot character” (Balcaen et al., 2006, P. 77). Business situations around companies change over time, and company failure may be not a sudden and unexpected event but a result of long-term processes and causes. Some characteristic behaviors and symptoms should be observed in a certain time frame (Luoma & Laitinen, 1991). Thus, it is important to consider the failure processes and figure out the specific pattern when predicting bankruptcy.

Chapter 3. Literature review: Evaluation system of classification accuracy

3.1 Introduction of evaluation system

Bankruptcy predictions are based on a binary classification: bankrupt or non-bankrupt. As should be the general approach to bankruptcy studies, the null hypothesis is that a company is not going to be bankrupted. Outcomes of the classification are generally shown in a following confusion matrix. The matrix is formed by four possible outcomes, which are True Negative, False Negative True Positive, and False Positive. The definitions of these terms are as follows:

1. True Negative (TN): negative prediction is correct. A company is predicted not to be bankrupted and the company is actually not bankrupted.
2. False Negative (FN): negative prediction is incorrect. A company is predicted not to be bankrupted but the company is actually bankrupted. It is also called as Type 2 error.
3. False Positive (FP): positive prediction is incorrect. A company is predicted to be bankrupted but the company is actually not bankrupted. It is also called as Type 1 error.
4. True Positive (TP): positive prediction is correct. A company is predicted to be bankrupted and the company is actually bankrupted.

Table 3-1 - Confusion matrix

Predicted / Actual	Non-Bankrupt	Bankrupt	Total
Non-Bankrupt	True Negative (TN)	False Positive (FP) (Type 1 error)	N
Bankrupt	False Negative (FN) (Type 2 error)	True Positive (TP)	P
Total	N*	P*	

(N = TN +TP, P = FN + TP, N* = TN + FN, P* = FP + TP)

With the confusion matrix, a number of useful measures to evaluate the bankruptcy prediction performance can be calculated. Especially, the following four measures are used in this study:

1. The rate of prediction accuracy shows how correctly the data were classified based on the prediction. This can be obtained as the total number of correct predictions (TP and TN) divided by the total number of observations.

$$\text{rate of prediction accuracy} = \frac{\text{TP+TN}}{\text{Number of observations in the data set}} \quad \text{(eqn.3)}$$

2. The error rate shows how much error was made when predicting bankruptcy and it is obtained as the total number of incorrect predictions (FP and FN) divided by the total number of the data set.

$$\text{error rate} = \frac{\text{FP+FN}}{\text{Number of observations in the data set}} \quad \text{(eqn.4)}$$

3. Sensitivity shows the true positive rate among the positive predictions. It is obtained as the number of correct positive predictions (TP) divided by the total number of positive actual outcome (P).

$$\text{Sensitivity} = \frac{TP}{P} \quad \text{(eqn.5)}$$

4. Precision show how correctly bankrupted companies are classified to be bankrupted. It is obtained as the number of correct positive predictions (TP) divided by the total number of positive predictions (P*)

$$\text{Precision} = \frac{TP}{P*} \quad \text{(eqn.6)}$$

3.2 Type 1 and Type 2 errors

When one model is claimed to be better than some other models, the common indicators are the Type 1 and Type 2 errors. As should be the general approach to bankruptcy studies, the null hypothesis is that a company is not going to be bankrupted and alternative hypothesis is that a company is going to be bankrupted. With respect to the hull hypothesis, the Type 1 and Type 2 errors are described as follows.

The Type 1 errors occur when the researcher rejects the null hypothesis, predicting that a company is going to be bankrupted, but in fact the company is not bankrupted. Using the confusion matrix, the Type 1 error rate is obtained by

$$\text{Type 1 error rate} = \frac{FP}{N} \qquad \text{(eqn.7)}$$

The Type 2 errors occur when the researcher fails to reject the null hypothesis, predicting that a company is not going to be bankrupted, but in fact the company is bankrupted. In bankruptcy prediction studies, making the Type 2 errors is more critical than making the Type 1 errors, causing higher costs and more severe consequences. Using the confusion matrix, the Type 2 error rate is obtained by

$$\text{Type 2 error rate} = \frac{FN}{P} \qquad \text{(eqn.8)}$$

Since two or more bankruptcy models are compared in bankruptcy prediction studies, it is inevitable to compare the outcomes to reach the conclusion. Previous studies confirmed the conclusion based on the Type 1 and Type 2 error rates. However, such studies focused on only the difference of the Type 1 and Type 2 error rates and did not focus on the differences of the contents of the Type 1 and Type 2 errors: which companies were actually misclassified. One company which is classified by one model is not always classified to the same group if a different model is used. Hence, it is arguable whether such an evaluation system yields robust conclusion regarding the claim of one model to be better than another model.

Chapter 4. Research Question

The principal aim of this study is to shed a light on the validity in claiming that one model is better than another model in the setting of bankruptcy studies. Historically within a scope of bankruptcy prediction, when a study claims that one model has “better” prediction ability, it means the Type 1 and Type 2 error rates of one model are lower than those of another model. For example, Begley et al. (1996) assessed the bankruptcy predictive ability of the Altman’s original Z-score (1968) and Ohlson’s logit model (1980) using datasets from 1980’s and concluded the outcomes based on the Type 1 and Type 2 error rates. They did not compare which companies are actually misclassified. This may cause problems to validate the research conclusions. The details are explained in the following examples.

4.1 Examples - Problems with the evaluation system

The following Matrices 1 to 4 show the pattern of classification errors which models may produce. In Matrix 1 and 2, there are four models (Model 1, 2, 3 and 4) in which the rate of prediction accuracy of correct classification is 80%. Matrix 1 shows that both Model 1 and Model 2 made errors on companies 4 and 5 when predicting failure. On the other hand, Matrix 2 shows Model 3 and Model 4 made errors on different companies: companies 2 and 8 and companies 4 and 5, respectively. Obtaining a result shown as in Matrix 1, a researcher must conclude that Model 1 and Model 2 have the same rate of the prediction accuracy and thus both models have same prediction ability (given the particular dataset). However, obtaining a result as shown in Matrix 2, the researcher also concludes that both models have the same prediction ability but the classification errors are made in different company. In such cases, it is not clear which of the models may be better. What should be concluded is that the models seem to be sensitive on some aspect of bankruptcy as contained within a particular sample and a particular set of independent variables.

Matrix 4-1 - Model 1 and Model 2

Company	1	2	3	4	5	6	7	8	9	10
Model 1	○	○	○	×	×	○	○	○	○	○
Model 2	○	○	○	×	×	○	○	○	○	○

(○=TN or TP, ×= Type 1 or 2 error)

Matrix 4-2 - Model 3 and Model 4

Company	1	2	3	4	5	6	7	8	9	10
Model 3	○	×	○	○	○	○	○	×	○	○
Model 4	○	○	○	×	×	○	○	○	○	○

(○=TN or TP, ×= Type 1or 2 error)

Another example is shown in Matrices 3 and 4. The rate of prediction accuracy of four models (Model 5, 6, 7 and 8) is different; in Models 5 and 7 it is 80% and in Models 6 and 8 it is 70%. In Matrix 3, Model 5 and 6 make the same errors on companies 4 and 5, and Model 6 makes one more error on company 6 when predicting failure. If the 10 sets of (independent) observations that yield the given classification pattern represented the entire population, only then can a researcher truly conclude that Model 5 is better than Model 6. In contrast, in Matrix 4, Model 7 and 8 show a different classification pattern. Model 7 made errors on companies 2 and 8 and Model 8 made errors on companies 4, 5 and 7. In this case, researchers cannot conclude that Model 7 predicts bankruptcy better than Model 8: if the models were to be applied a dataset which has 5 more companies, the pattern of the new 5 classifications will potentially invalidate the earlier conclusion about which model performed better.

Matrix 4-3 - Model 5 and Model 6

Company	1	2	3	4	5	6	7	8	9	10
Model 5	○	○	○	×	×	○	○	○	○	○
Model 6	○	○	○	×	×	×	○	○	○	○

(○=TN or TP, ×= Type 1or 2 error)

Matrix 4-4 - Model 7 and Model 8

Company	1	2	3	4	5	6	7	8	9	10
Model 7	○	×	○	○	○	○	○	×	○	○
Model 8	○	○	○	×	×	○	×	○	○	○

(○=TN or TP, ×= Type 1or 2 error)

Based on the apparent insufficiency to truly determine which model may be superior in a bankruptcy setting, my research question is:

Do the reported results of the MDA and Logit models in the literature with respect to bankruptcy prediction hold when firm-specific classification patterns are considered?

Chapter 5. Research methodology and data

5.1 Econometric models

There is no generally accepted theory of bankruptcy but of course there is a theory which is associated with any particular (statistical) method. In this research, two econometric models, the MDA and Logit models, are used to classify firms. The following section is a short summary of the two models.

5.1.1 MDA

MDA is a traditional statistical method used to classify an observation into one of several groups depending on the character of the individual observations (Altman, 2000). In the case of bankruptcy analysis, it is used for classifications where the response variable is binary, such as bankrupt or non-bankrupt, and making predictions eventually. MDA can be expressed through a discriminant function as follows:

$$Y = d_0 + d_1X_1 + d_2X_2 + \cdots + d_nX_n + e \qquad \text{(eqn.9)}$$

where Y is the discriminant score; d_0 is the intercept, d_j is the discriminant coefficient (j = 1,2,…,n), X_j are the independent variables, and e is a random error. The training dataset is used to derive an estimated linear combination based on the characteristics of factors and coefficients are estimated.

The discriminant function (eqn.9) combines the number of different independent variables into one single multivariate discriminant score, $Y \in (-\infty, \infty)$, which indicates the financial healthiness of companies as follows: lower Y indicate poorer financial healthiness. Based on the discriminant score Y and some cut-off ranges, companies are classified (predicted) to the bankrupt or non-bankrupt group. The cut-off point is determined in advance and used as a threshold of the MDA model. When the discriminant scores are lower than the cut-off point, they are classified as a bankrupt group entity, whereas when the discriminant scores are higher than the cut-off point, they are classified as a non-bankrupt group entity.

MDA is based on several assumptions. E.g., it assumes (in approximation) dichotomous data, Normality of the error term, equal dispersion of variance-covariance matrices across two groups, and the absence of multicollinearity (Balcaen and Ooghe, 2006).

5.1.2 The Logit model

Regression models can be used to make predictions based on the relationship between independent variables and dependent variables. A logit model is one type of regression model, and it is particularly useful to find possibilities to which group the response is classified. The notable advantage using the logit model is to overcome the problem that the true value of outcome probabilities may be greater than 1 or smaller than 0. The logit model, removes the upper bound by using odds, and removes the lower bound by using the logarithm. Hence, the logit model is more suitable for interpreting the probabilities with a qualitative response. The logit model is described as follows:

If P is probability of an event and D is the odds of the event,

$$D = \frac{P}{P-1} = \frac{Probability\ of\ event}{Probability\ of\ no\ event} \ , \qquad \text{(eqn.10)}$$

and assuming that the logit value of the probability is a linear function

$$\log\left[\frac{P(x)}{1-P(x)}\right] = \beta_0 + \beta_1 X_1 \quad 0 \le P(x) \le 1 \ , \qquad \text{(eqn.11)}$$

One can solve the equation

$$P(x) = \frac{e^{\beta 0+ \beta 1\, X 1}}{1+ e^{\beta 0+ \beta 1\, X 1}} \qquad \text{(eqn.12)}$$

And simplify to the logit function form as follows:

$$P(x) = \frac{1}{1+ e^{-\beta 0- \beta 1\, X 1}} = \frac{1}{1+ e^{-z}} \text{ , (where z = } \beta_0 + \beta_1\, X_1) \qquad \text{(eqn.13)}$$

The dependent variable of the logit model can be any value, as long as it is consistent with $0 \le P(x) \le 1$.

The logit model is employed by many researchers in predicting bankruptcies as the model suits particularly well to capture the nature of bankruptcy prediction (Ohlson, 1980, Back et al., 1996, Ho et al., 2013). The response variables in bankruptcy prediction are qualitative and in most cases they are binary; bankrupt and non-bankrupt. Also the logit model can be used when more than two response variables are required, such as classifying a company as bankrupt, non-bankrupt, and might-be-bankrupt.

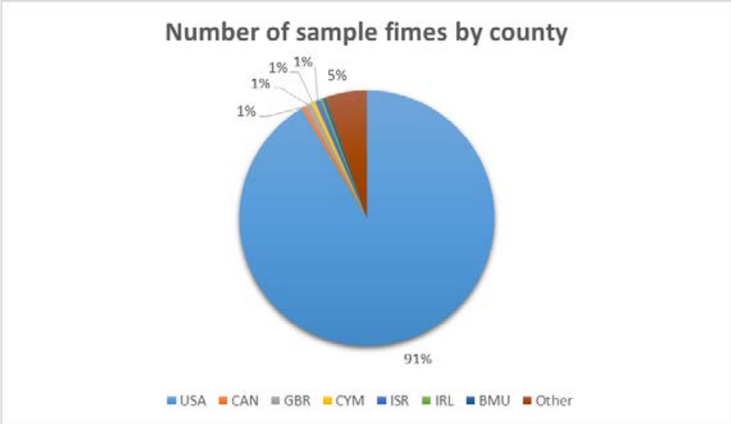
5.2 Data

5.2.1 Data source

The initial sample of this study is obtained from the Compustat database (through CRSP). The database contains financial, statistical and market information for active and inactive publicly traded US-listed companies. Ohlson (1980) obtained a sample of non-bankrupt firms from Compustat of 2,058 for his period of analysis (1970-1976). The database can provide not only annual and quarterly financial information but also non-financial information such as countries, industry classifications, a size of the companies, IPO dates, and delisted dates and the reasons for delisting. The database contains this information also for non-US firms that list in one of the US markets.

Table 5-1 shows the firm year distribution with respect to the country where the company headquarters are situated. The dataset is extracted from the Compustat database to understand the contents of the database and create appropriate samples for the research. The dataset includes bankrupt and non-bankrupt companies from all over the world from 1964 to 2012. It contains 320,088 firm-years for companies from 63 different countries, of which the number of US firm-years is 292,696.

Table 5-1 – Number of extracted sample firms by country (USA:USA, CAN: Canada, GBR: United Kingdom, ISR: Israel, IRL: CYM: Cayman Islands, BMU: Bermuda)



Generally, a larger sample is expected to reflect the actual composition of firms on the US equity market better. In other words, the accuracy of coefficient estimates and the power of a model predicting bankruptcy are expected to increase while the number of firms in the sample increases. Therefore, US companies are the principal target in this research and companies outside US are excluded because of the comparability of the economic context.

5.2.2 Delisted companies

The Compustat database supplies specific information about delisted companies, such as a delisting date and delisting reason. There are 9 reasons the database supplies. Table 5-2 illustrates the number of U.S. bankrupt companies classified into the 9 delisted reasons from 1964 to 2012.

Table 5-2 – The number of US bankruptcy companies by delisted reason

dlrsn	Delisted reason	Number of firms
1	Acquisition or merger	10006
2	Bankruptcy	960
3	Liquidation	849
4	Reverse acquisition (1983 forward)	119
5	No longer fits original format (1978 forward)	77
6	Leveraged buyout (1982 forward)	88
7	Other (no longer files with SEC among other possible reasons), but pricing continues	983
9	Now a private company	501
10	Other (no longer files with SEC among other reasons)	2921

<u>Bankruptcy</u>

Bankruptcy is one of the major exits for distressed firms. In the Compustat database, companies are deleted under bankruptcy when a company files for reorganization under Chapter 11 of the U.S. bankruptcy law. Chapter 11 is defined as a reorganization bankruptcy because it involves restructures of debtor's business affairs. A company that files under Chapter 11 generally faces unmanageable debt problems and needs (obtains) time to restructure their financial situation. The company is expected to turnaround their business and become profitable again. During being in Chapter 11, the firm is subject to fulfilling obligations under the plan of reorganization. This type of reorganization is generally complex and expensive. There are many U.S. organisations who have filed Chapter 11 to reorganize their business such as General Motors, United Airlines, and K-mart. The major reason for seeking Chapter 11 protection is that it allows distressed firms to continue to run their business. If the reorganization process fails, the distressed company enters the liquidation process: all assets are liquidated and paid off to stakeholders.

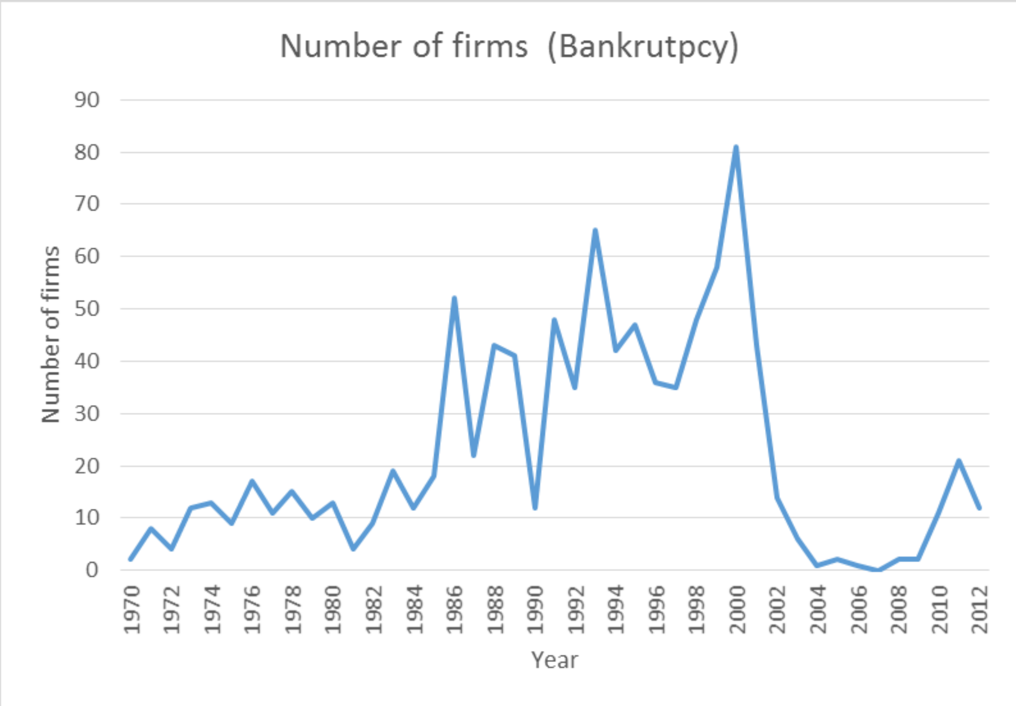
The data of bankrupt firms are available from 1970 onwards. The total number of bankrupt firms from 1970 to 2012 is 960. The number starts to increase from the 1980ies, then hits a peak of 81 bankruptcies recorded in 2000 (Graph 5-1). The major reason for the increase of bankruptcy in 1980ies is due to a number of developments of bankruptcy rules from late 1970ies to 1980ies.[3] The most notable change is the Bankruptcy Reform Act of 1978 which took effect on October 1, 1979, in which business reorganization Chapter, such as Chapter11 has been introduced.[4]

---

[3] The securities industry saw significant turbulence in 1969 and 1970, leading to voluntary liquidations, mergers, receiverships and bankruptcies of a substantial number of brokerage houses. In reaction to this situation, Congress enacted the Securities Investor Protection Act of 1970 in an attempt to quell the filings, restore investor confidence and upgrade financial responsibility requirements for registered brokers and dealers.

[4] The old Chapters X, XI was replaced to Chapter 11.

Graph 5-1 - Number of bankruptcy firms

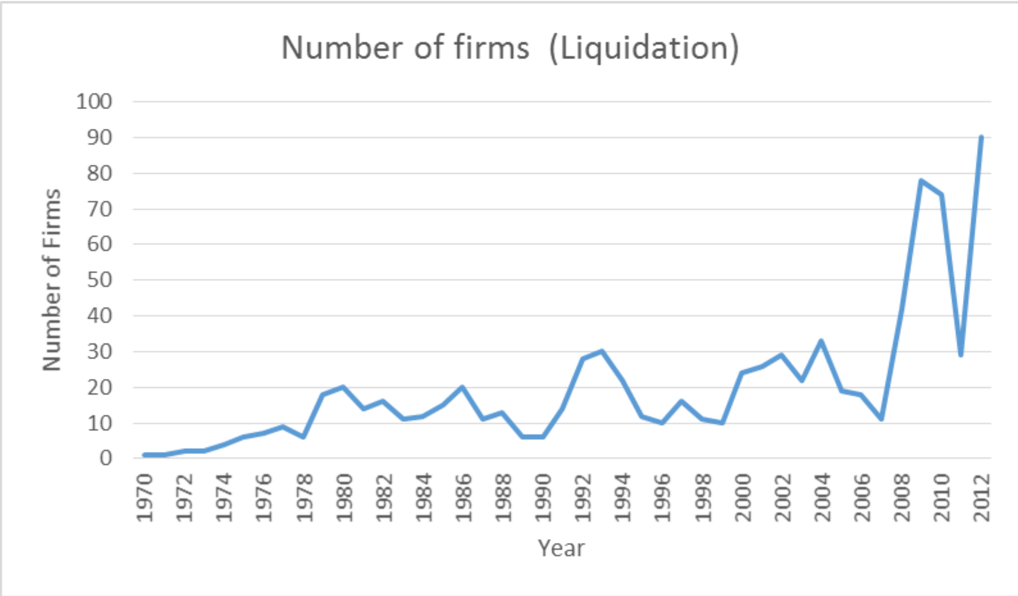


<u>Liquidation</u>

Liquidation is another major exit for distressed firms. In the Compustat database, companies are deleted under bankruptcy when a company files for reorganization under Chapter 7 of the U.S. bankruptcy law. Under Chapter 7, the company stops all operations and goes completely out of business. A trustee is appointed to liquidate (sell) the company's assets, and the money is used to pay off debt. The investors and creditors are paid off by absolute priority rule which stipulates the order of payment: creditors before shareholders. Secured creditors are paid off first because the credit is normally tied to a collateral. The second line is unsecured creditors, the third line is bondholders, and lastly shareholders are paid off.

The data of liquidation firms is available from 1970 onwards. Graph 2 shows the number of liquidations filed in the US from 1970 to 2012. It is interesting to see how the trend of liquidation is different from bankruptcy: the number of liquidation firms has some fluctuations until 2007 when there was a sudden increase from 2008 to 2009. It jumped up to 78 in 2009 from 11 in 2007, but fell down to 29 in 2011 and raised strongly again in 2010. The highest number is recorded with 90 liquidations in 2012. Total number of liquidation from 1970 to 2012 is 849.

Graph 5-2 - The number of liquidation firms



<u>Bankruptcy – Trading companies' shares and public disclosure</u>

As far as Chapter 11 allows the distressed firms to operate their business as normal, the firms might be able to continue to publicly trade after the filing date. There is no federal law that prohibits trading of shares of such companies. In most instances, however, companies that file under Chapter 11 are unable to meet the requirements to continue to trade set by the security offices of major stock exchanges in the US, such as, NASDAQ or the NYSE. There are still chances that their shares may continue to trade on either the OTCBB or the Pink Sheets.

Since the operation of the distressed companies continues, the company information can be publicly disclosed. However, preparing periodic reports, particularly audited financial statements, can be expensive as it involves substantial accounting and legal expenses. Also, maintaining compliance with their Exchange Act reporting obligations is difficult for most distressed firms. Therefore, the companies determine whether to continue the periodic reporting comparing the benefit and the cost. In other words, the availability of accounting and finance information of the distressed firms depends on the decision of companies. All of this is relevant for deciding about what data ought to be used in a bankruptcy study.

5.3 Sample

This analysis is based on the research by Begley et al. (1996). It is because the research is to assess the accuracy of prediction two models: MDA (Altman, 1968) and Logit (Ohlson, 1980) models. Begley et al. used datasets from the 1980ies and re-estimated the two models. Their results show that the Type 1 and Type 2 error rates of the re-estimated models are lower than the originals' (Altman, 1968; Ohlson, 1980) and, therefore, they conclude that the re-estimated models calculated by the new datasets have more accurate prediction power than the original models.

In this section, firstly I analyse the dataset and create a basic sample. Based on the basic sample, I create new samples replicating Begley's results so as to examine their research conclusions.

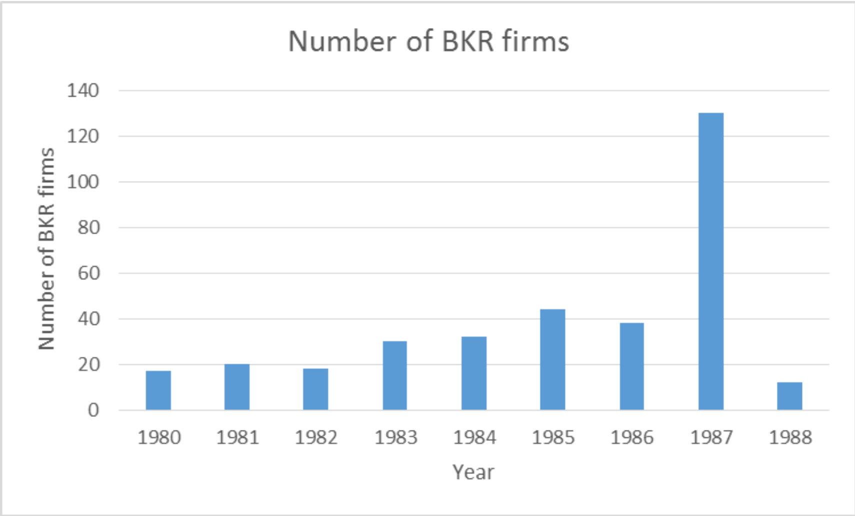
5.3.1 Data preparation

In order to create a basic sample from which I can later reproduce Begley et al.'s results and draw my own sample, the following criteria must hold:

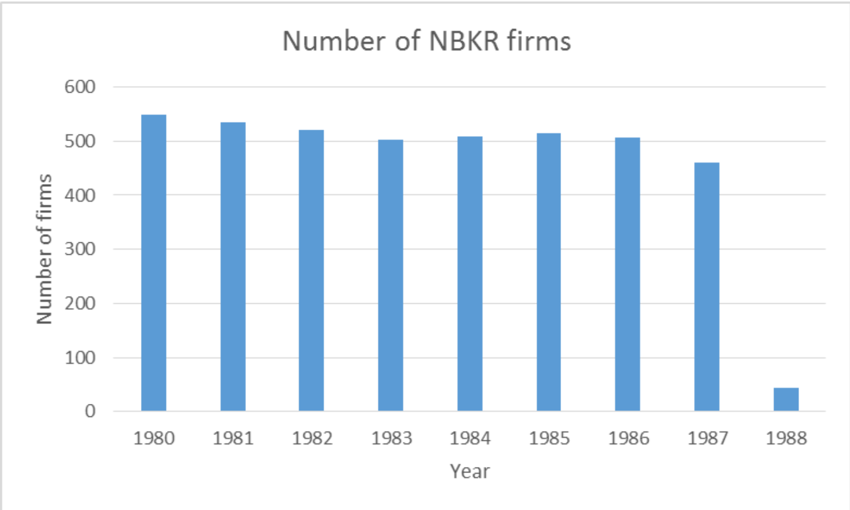
1) Must be a US firm;
2) The firms are listed (share prices are available in Compustat);
3) There is sufficient financial statement information for more than two years to estimate both Altman's and Ohlson's models;
4) All bankrupted firms (cf. Table 5-2) but those with dlrsn=2 (Chapter 11 bankruptcy) are removed.

Based on above 4 rules, the following sample haven obtained (Graph 5-3 and Graph 5-4).

Graph 5-3 - Number of bankrupt firms



Graph 5-4 - Number of non-bankrupt firms



5.3.2 Replicate the Begley et al. sample – bankruptcy firms

Begley et al. (1996, p.4) created their sample of bankrupt firms as follows:

1) Financial statement data are available from Compustat during the time period from1980 to 1989 and is sufficient to estimate the Altman and Ohlson models;
2) The firm is listed and there is an observable market price at the end of the fiscal period for which the financial (accounting) statement information is being used;
3) SIC code is smaller than 4000 or between 5000 and 5999;

4) The firm had a bankruptcy footnote on Compustat or had a dlrns=2 reason of default;
5) The firm must not have filed for bankruptcy within the previous five years;
6) The last available financial statement prior to bankruptcy must be for a Compustat year falling between 1980 and 1989;
7) The firm's total assets must exceed $10 million;
8) The bankruptcy filing date must follow the fiscal year end by at least four months. If the bankruptcy filing occurs within four months following the fiscal year end, the previous year's financial statements are treated as the last financial statements available prior to bankruptcy; and
9) The time lag between the fiscal year end of the last available financial statements prior to bankruptcy and the bankruptcy filing date must be no more than 18 months.

The final sample of bankrupt firms in Begley et al.'s paper is 165. From this sample 100 bankrupt firms were randomly selected as a training dataset, and used for estimating the two bankruptcy prediction models. The remaining 65 firms were retained as a hold-out sample to test the two models. In order to compare the results of predictive ability of two models with Begley et al.'s research, I replicated the samples as best possible.

Using the basic sample created earlier, I then have adopted the first six criteria listed above without any change. Some of the other criteria have been changed which is indicated below including a rational for such change.

7') *'The firm's total assets must exceed $10 million.'* I limited the selection of firm's total assets must be less than $200 million. This restriction is created not to eliminate outliers a priori but to not have unnecessarily unrepresentative data skew the parameter estimates.

8') *'The bankruptcy filing date must follow the fiscal year end by at least four months. If the bankruptcy filing occurs within four months following the fiscal year end, the* *previous year's financial statements are treated as the last financial statements available prior to bankruptcy.'* → Not adopted. The deletion date given in the Compustat database is frequently years after the last financial information entries.

9') *'The time lag between the fiscal year end of the last available financial statements prior to bankruptcy and the bankruptcy filing date must be no more than 18 months.'* → Not adopted. The bankruptcy process evolves over long periods of time (years). To arbitrarily determine cut-off durations creates bias.

Limitation of replicating the Begley et al. sample

1) Chapter 11 filing date

Begley et al. used the Compustat database. Firstly, they went through criteria 1) to 4) above. This initial sample was 466. Further to that, they obtained Chapter 11 filing dates from the *Capital Changes Reporter, Lexis/Nexis, The directory of Obsolete Securities* and other researchers. They obtained Chapter 11 filing dates for 396 out of 466 bankrupt firms.

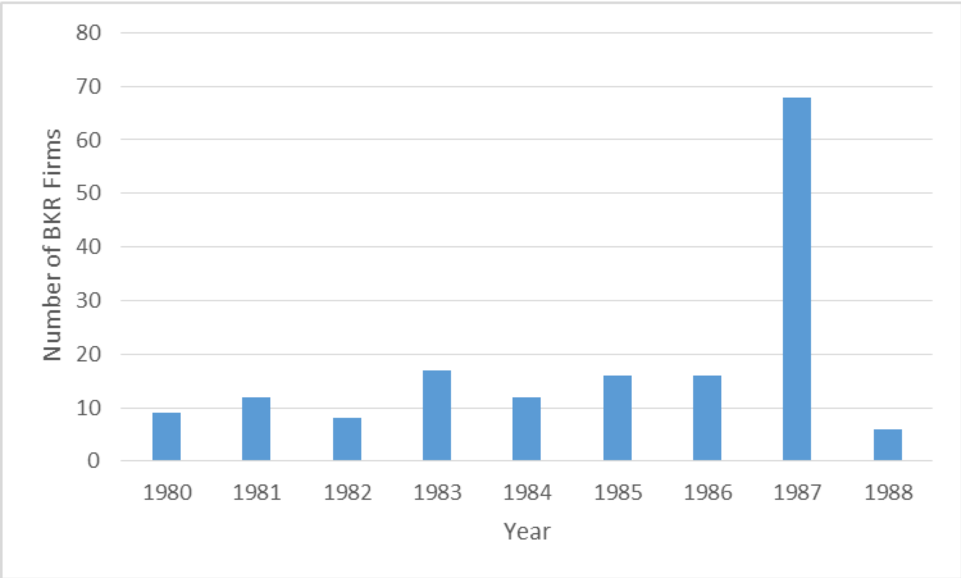
During my research, I attempted to obtain Chapter 11 filing dates but reaching reliable data sources is difficult: I compared the time lag between Compustat delisted dates and last financial statements which supply sufficient accounting information to estimate both Altman's and Ohlson's models. The results showed there are only a few bankrupt firms which are delisted in less than two years after the last available financial statement. From this it is obvious that the sample number is not adequate to estimate the two models, and therefore, this research is not following the 5th and 6th criteria in Step 2.

2) Data source

The Compustat database is available at University of Canterbury, but the version available does not supply the financial accounting data necessary to calculate financial ratios to estimate the two models. Therefore, the dataset of this research has been sourced elsewhere.

After following the criteria 1) to 6) and 7'), I obtained a sample of 164 bankrupt firms (cf. Graph 5-5).

Graph 5-5 Number of bankruptcy firms



From all 164 bankruptcy firms, 100 bankrupt firms are randomly selected to re-estimate the two models. 64 remaining bankrupt firms are used to create a hold-out sample for testing the predictive ability of the re-estimated models.

5.3.3 Replicate the sample – non-bankruptcy firms

Begley et al. selected the sample of non-bankrupt firms as follows:

1) Compustat does not classify the firms as bankrupt firms during the previous five years; and
2) The firms do not bankrupt within the following two-and -half years.

They created two different samples of non-bankrupt firms to keep consistency with the different requirements of Altman and Ohlson's estimation processes.

<u>Sample to re-estimate Altman's model</u>

The sample of non-bankrupt firms to match the 100 bankrupt firms is a matched pair sample according to year, SIC code, and firm size. The ratio of bankrupt firms and non-bankrupt firms is thus 1:1.

Following Begley et al.'s modified criteria, I replicated a paired sample of non-bankrupt firms as follows:

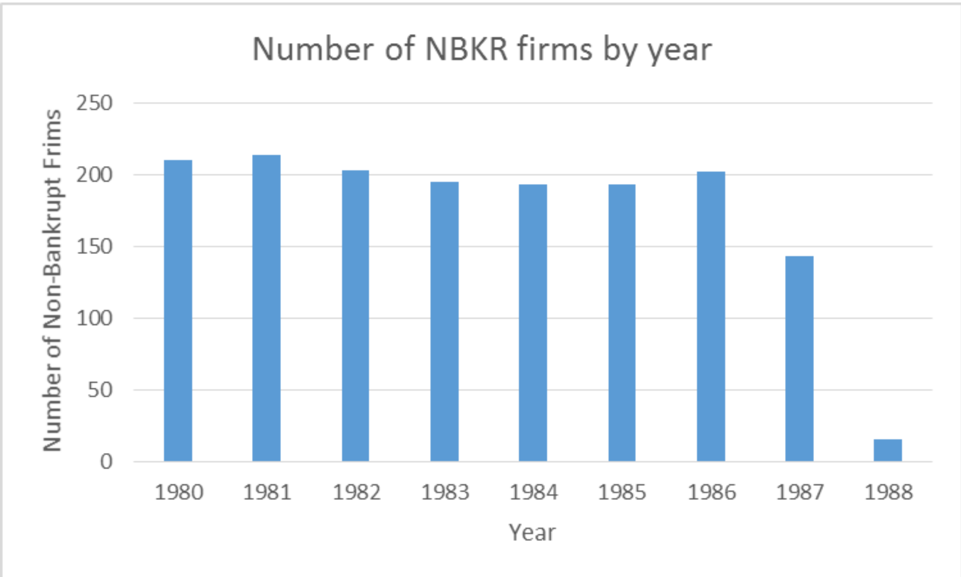
1) Compustat year from 1980 to 1988;

2) SIC code is less than 4000 or between 5000 and 5999, and

3) Total asset must exceed $10 million but less than $200 million.

After applying all three criteria, 366 distinctive non-bankrupt firms are obtained which provide 1569 firm year observations, as shown in Graph 5-6. From this sample, a paired sample of non-bankrupt firms of 164 is created. Non-bankrupt firms are selected from each year in order to match the number of bankrupt firms in each year. The ratio of bankrupt firms and non-bankrupt firms is thus also 1:1.

Graph 5-6 - Number of non-bankrupt firms



Sample to re-estimate Ohlson's model

Begley et al. created a sample of non-bankrupt firms by randomly selecting 200 non-bankrupt firms from each year between 1980 and 1989. Sample firms are selected only once. Total sample of non-bankrupt firms is 2’000, and the ratio of bankrupt firm versus non-bankrupt firms is thus 1:20.

The matching proportion of bankrupt and non-bankrupt firms of 1:20 cannot be obtained from my data - the basic sample does not contain a sufficiently high number of non-bankrupt firms to create the sample. In order to keep the matching proportion of 1:20, the condition of creating a paired sample is removed and non-bankrupt firms can be selected more than once.

Chapter 6. Empirical work analysis

6.1.1 Re-estimated models

MDA – Altman's model

My sample consists of 164 bankrupt firms and 164 non-bankrupt firms. Total number of sample firms is 328. The yearly distribution of firms is shown below:

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of BKR Firms	9	12	8	17	12	16	16	68	6	164
Number of NBKR Firms	9	12	8	17	12	16	16	68	6	164

From this 328 firms were split into a training data set and a testing dataset. A training dataset includes 100 bankrupt firms and 100 non-bankrupt firms. A remaining 64 bankrupt firms and 64 non-bankrupt firms are used for testing dataset.

Variables and preliminary data analysis

Altman (1968) used multiple discriminant analysis (MDA) with five variables that showed strong discriminating power. The five variables are:

WC = Working capital / Total assets

RE = Retained earnings / Total assets

EBIT = Earnings before interest & taxes / Total assets

ME = Market value equity / Book value of total debt

SA = Sales / Total assets

Group means of each variable and *p*-values for a two-sided mean difference Z-test are shown below:

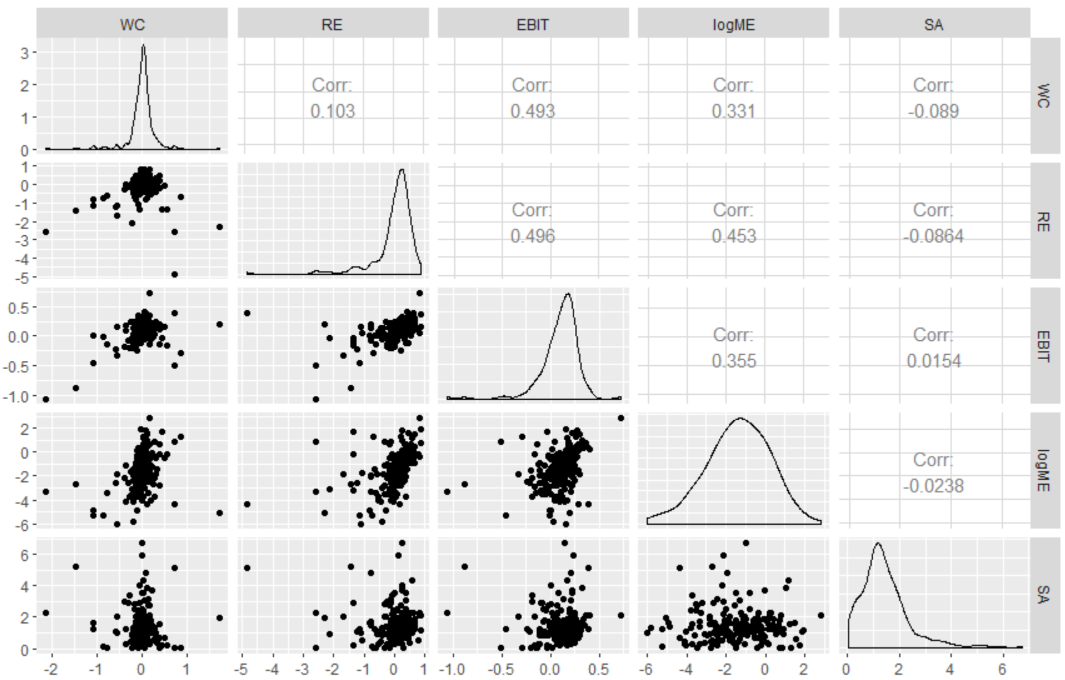
```
                     Group means for variables
             WC          RE        EBIT    Log(ME)        SA
0    0.05861546   0.2499771  0.16769530   -0.63827  1.441013
1   -0.05879417  -0.2534957  0.03409139   -2.15268  1.484594
p        0.0084     7.1E-09    1.05E-07   3.11E-14     0.770
(0 = non-bankrupt firms, 1 = bankrupt firms)
```

The mean values of the first four variables for bankrupt firms are significantly smaller, at the 5%-level, than those for non-bankrupt firms. This means each variable has the power to discriminate the firms into two groups. The fifth variable SA has means that are not significantly apart for bankrupt and non-bankrupt firms – a finding consistent with Altman’s (1968) work.

Correlation between the five variables

Table 6-2 shows that the relationship between the five variables. The positive correlation between WC, EBIT and RE is not surprising in that all three numerators are profit measures of the company. Note that I have transformed ME to logME because the latter approximates the Normal much better than the raw ME data - at the cost of higher correlations with the three profit-related variables WC, EBIT and RE. These positive relationships are consistent with findings from the capital markets literature where earnings and market price evolve dynamically within bounds of on another (e.g., Falta & Willett, 2013).

Table 6-1 - Correlation table for Altman's five variables



Model fitting

After ascertaining the power of 4 of the 5 variables to categorise, on average, the dichotomous fail and non-fail data, the MDA fitting algorithm lda() in R produces the following coefficient estimates that best discriminate the data into the two groups:

Re-estimated coefficients by MDA[5]

Coefficients of linear discriminants:	
	LD1 = 'order 1 linear discriminant system'
WC	0.26948297
RE	-0.43518896
EBIT	-2.09790946
logME	-0.50834666
SA	0.01494353

The re-estimated model thus is:

$$\boldsymbol{Z = 0.269X_1 - 0.435X_2 - 2.097X_3 - 0.508X_4 + 0.014X_5} \qquad \text{(eqn. 14)}$$

<u>Logit – Ohlson's model</u>

Sample

To re-estimate the Ohlson model, the same sample of bankrupt firms is selected as for the MDA reestimation. I thus partially limit the inconsistency of sample bias. However, the paired non-bankrupt firms are randomly selected from the basic sample. As the matching proportion of bankrupt and non-bankrupt firms for Ohlson's model is one to twenty, training dataset consists 100 bankrupt firms and 2000 non-bankrupt firm (years).

Variables

Ohlson's model used the following nine variables:

1. SIZE – In(total assets/GNP price level index). The index assumes a base value of 100 for 1968.
2. TLTA – Total liabilities/Total assets
3. WCTA – Working capital/ Total assets

---

[5] Compared to the original Z-score, the coefficients of re-estimated model are negative. It is because the re-estimated models used Y=1 for the bankrupt firms (Altman et al., 2017).

4. OLCA – Current liabilities/Current assets
5. OENEG – One if total liabilities exceed total assets, otherwise zero
6. NITA – Net income/Total assets
7. FUTL – Funds provided by operations/Total liabilities
8. INTWO – One if net income was negative for the last two years, otherwise zero
9. CNHIN – $(NI_t – NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$, where NI_t is net income for the most recent period

Group means of each variable and *p*-values for a two-sided mean difference Z-test are shown below:

	SIZE	TLTA	WCTA	CLCA	OENEG
0	-1.317833	4.384880	0.02455292	0.6637502	0.014
1	-2.024185	2.011359	-0.05879417	1.3309139	0.140
p	>0.001	0.0019	0.051	0.0618	n/a

	NITA	FUTL	INTWO	CHIN
0	0.02176659	0.9381050	0.118	0.02083205
1	-0.14836618	-0.1488101	0.110	-0.24426401
p	>0.001	0.1468	n/a	>0.001

0 = bankrupt firms, 1 = non-bankrupt firms

On the univariate level four variables (SIZE, TLTA, NITA, CHIN) display discriminatory power, and three (WCTA, CLCA, FUTL) do not, at the 5% significance level.

Model fitting

Using R's glm() package yields the following coefficient estimates, the majority of which are statistically significant at the 5%-level:

```
Deviance Residuals:
     Min          1Q     Median        3Q        Max
-1.6693      -0.3094   -0.2506      -0.2004    3.4590

Coefficients:
                Estimate  Std.   Error     z value   Pr(>|z|)
(Intercept)     -2.66951         0.25347   -10.532   < 2e-16  ***
SIZE            -0.18122         0.06058   -2.991    0.002778  **
TLTA            -0.27472         0.09761   -2.814    0.004885  **
WCTA            0.29054          0.38495    0.755    0.450409
CLCA             0.03045         0.02587    1.177    0.239159
OENEG           1.62556          0.43251    3.758    0.000171  ***
NITA            -0.94472         0.59521   -1.587    0.112466
FUTL            -0.28710         0.15512   -1.851    0.064194 .
INTWO           -1.55251         0.46846   -3.314    0.000919  ***
CHIN            -0.65736         0.19694   -3.338    0.000844  ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 804.07  on 2099  degrees of freedom
Residual deviance: 705.64  on 2090  degrees of freedom
AIC: 725.64
```

```
Number of Fisher Scoring iterations: 10
```

I have further checked if presence of multicollinearity would bias the coefficient estimates. Using R's vif() routine, TLTA, NITA and FUTL show signs of introducing potential bias through using the same denominator. However considering the on-going discussion around VIF (variance inflation factor) cut-off points, values of between 5 and 10 may imply multicollinearity and values below 5 are deemed to have a system of negligible to no impact due to multicollinearity (Kim & Kang, 2012). Overall I thus conclude that multicollinearity is of little relevance in the logit model.

	SIZE	TLTA	WCTA	OENEG	NITA	FUTL	INTWO	CHIN
vif	1.169	4.126	2.765	1.767	4.639	5.010	1.105	1.289

6.1.2 Cut-off points

The cut-off points are scores (MDA) or a probability (Logit) to classify the firms into bankrupt and non-bankrupt groups. At the optimal cut-off point, the total misclassification error is minimized. Altman (1968) found that the optimal cut-off point for the original Z-score was 2.675 and Begley et al. (1996) adopted the same optimal cut-off point for their re-estimated MDA model. Ohlson (1980) found that the optimal cut-off point for the Logit model was 0.038 while Begley et al. found a different optimal cut-off point of 0.061 for the re-estimated model. The choice of optimal cut-off points made a significant impact on the prediction accuracy of the models as it directly influences the Type 1 and Type 2 error rates. Bearing in mind the lack of any theoretical rational researchers may use to determine a cut-off point, the choice indeed is based on outcome driven optimization.

In this research, I use the Receiver Operating Characteristic (ROC) curve to find out the optimal cut-off point for the Logit model. ROC curves take into account all possible cut-off points and display the corresponding Type 1 and Type 2 errors. They are useful not only for comparing different classifiers but also to find out the optimal cut-off point. Appendix 1 displays the ROC curves, a reference to a description on how to determine the cut-off points, and the numerical values of the cut-off points used for further analysis.

Given the estimated models and determined cut-off points, in the remaining section of this chapter, I am going to assess the predictive ability of the MDA and Logit models under various criteria.

6.1.3 Evaluation of predictive ability

Testing sample

Two testing samples are created. One is for testing the predictive ability of the MDA model and another is for testing a predictive ability of the Logit model. The testing sample contains the 64 out-of-estimation-sample firms reserved (cf. Chapter 5), which have been randomly matched 1:1 with non-bankrupt firms. The final testing sample thus contains 128 companies which need be classified into two groups, bankrupt (BKR) and non-bankrupt (NBKR). Matrices 6-1 and 6-2 show that the MDA model commits fewer classification errors (33/128=26%) than the Logit model (37%).

Matrix 6-1 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	55	9
	BKR	24	40
Prediction accuracy		74%	
Type 1 error (FP)		14%	
Type 2 error (FN)		38%	

Matrix 6-2 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	53	11
	BKR	36	28
Prediction accuracy		63%	
Type 1 error (FP)		17%	
Type 2 error (FN)		56%	

However, reducing the testing sample to a random selection of 10 out of the 64 reserved bankrupt firms and match these with drawing randomly 10 non-bankrupt firms from the non-bankrupt holdout sample, the above result is reversed and the logit model comes out on top with respect to forecasting ability (cf. Matrices 6-3 and 6-4).

Matrix 6-3 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	6	4
Prediction accuracy		60%	
Type 1 error		20%	
Type 2 error		60%	

Matrix 6-4 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	9	1
	BKR	4	6
Prediction accuracy		75%	
Type 1 error (FP)		10%	
Type 2 error (FN)		40%	

While both testing samples contain an arbitrary number of companies, be it the originally reserved 64 bankrupt firms or the reduced testing sample containing 10 bankrupt firms, the different predictive model performances are sample-driven results.

Matrices 6-5, 6-6 and 6-7 show the classification decision at the company level. These matrices directly address the research question.

Matrix 6-6 Error pattern of MDA and Logit models (128 firm testing sample)

MDA	Firm #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Predicted	0	1	1	1	1	1	0	0	1	0	1	1	1	0	0	0	0	1	0	1
	Actual	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	FP	TP	TP	TP	TP	TP	FP	FP	TP	FP	TP	TP	TP	FP	FP	FP	FP	TP	FP	TP
Logit	Predicted	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	1	1	0
	Actual	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	FP	FP	FP	FP	TP	TP	FP	FP	FP	FP	TP	TP	FP	FP	FP	TP	TP	TP	TP	FP

MDA	Firm #	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	Predicted	0	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	1
	Actual	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	FP	FP	TP	TP	TP	TP	TP	TP	TP	FP	FP	TP	TP	TP	TP	TP	FP	FP	FP	TP
Logit	Predicted	0	0	1	1	1	0	1	1	0	0	0	1	1	0	1	1	0	0	0	0
	Actual	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	FP	FP	TP	TP	TP	FP	TP	TP	FP	FP	FP	TP	TP	FP	TP	TP	FP	FP	FP	FP

MDA	Firm #	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
	Predicted	1	0	1	1	0	0	1	1	1	1	0	1	1	1	1	0	1	0	0	1
	Actual	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	TP	FP	TP	TP	FP	FP	TP	TP	TP	TP	FP	TP	TP	TP	TP	FP	TP	FP	FP	TP
Logit	Predicted	1	0	0	0	0	0	1	0	1	1	1	1	0	0	1	0	0	0	0	1
	Actual	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	Class	TP	FP	FP	FP	FP	FP	TP	FP	TP	TP	TP	TP	FP	FP	TP	FP	FP	FP	FP	TP

MDA	Firm #	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
	Predicted	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0
	Actual	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	TP	TP	TP	FP	TN	TN	TN	TN	TN	FN	TN	TN	TN	FN	TN	FN	TN	TN	TN	TN
Logit	Predicted	1	1	1	0	0	0	0	0	1	1	0	0	0	1	0	1	1	0	0	0
	Actual	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	TP	TP	TP	FP	TN	TN	TN	TN	FN	FN	TN	TN	TN	FN	TN	FN	FN	TN	TN	TN

MDA	Firm #	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
	Predicted	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
	Actual	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	FN	TN	TN
Logit	Predicted	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
	Actual	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	FN	TN	TN	TN	TN	TN

MDA	Firm #	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
	Predicted	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
	Actual	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TN	TN
Logit	Predicted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	Actual	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Class	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN

MDA	Firm #	121	122	123	124	125	126	127	128
	Predicted	1	0	0	0	0	0	0	1
	Actual	0	0	0	0	0	0	0	0
	Class	FN	TN	TN	TN	TN	TN	TN	FN
Logit	Predicted	121	122	123	124	125	126	127	128
	Actual	0	1	1	0	0	0	0	0
	Class	0	0	0	0	0	0	0	0
		TN	FN	FN	TN	TN	TN	TN	TN

Matrix 6-6 Error pattern of MDA model (20 firm testing sample)

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	0	1	0	0	1	0	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	TP	FP	FP	TP	FP	FP	TP	FP	FP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	0	0	1	0	0	1	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	TN	FN	TN	TN	FN	TN	TN

Matrix 6-7 Error pattern of Logit model (20 firm testing sample)

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	0	1	1	1	1	1	0	0
Actual	1	1	1	1	1	1	1	1	1	1
Class	FP	FP	FP	FP	TP	TP	TP	TP	TP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	1	0	0	0	0	0	0	0
Actual	0	0	1	0	0	0	0	0	0	0
Class	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN

I first discuss the results for the 20 firm testing sample. Let $N_{common} = N_c$ be the number of firms wronlgy categorised by both models (N_c =4; firms 2, 3, 9 and 15), and $N_{separate} = N_s = N_{MDA} + N_{Logit}$ be the numbers of firms wronlgy categorised by either model ($N_s = 5$; $N_{MDA} = 4$, firms 5, 6, 8 and 18; $N_{Logit} = 1$, firm 10). Let N (= 2 x 20; two samples with 20 observations each) be the total number of classification mistakes possible, and N_{agg} (=13) be the aggregate number of mistakes committed by both models. In case of the 128 firm testing sample, N_c =25, $N_s = 30$, $N_{MDA} = 10$, $N_{Logit} = 20$, N = 256 and $N_{agg} = 80$.

The following qualitative assessments can be made. For example, if one, but not the other of the two N_{MDA} and N_{Logit} equals zero we obtain the situation depicted in Matrix 4-3, which is a case where the claim 'one model be superior to another model' can be better substantiated than when both N_{MDA} and N_{Logit} are not equal zero and the ratio between common mistakes and individual mistakes on parts of either model are large (which indicates that both models pick up different signals and failure modes). Given that this is the case in my analysis for both testing samples sets where both N_{MDA} and N_{Logit} are large in comparison to N_C and N_{agg}, none of the two models can be deemed to have superior predictive ability.

The next Chapter addresses the differences in the mixing ratios of the two samples which were used to estimate the model coefficients and the difference in the variable selection process, which seem to produce incomparable conditions to truly establish model superiority in bankruptcy analysis.

Chapter 7. A closer look at sample composition and variable selection

In Chapter 7, I analyse the sensitivity of the Type 1 and Type 2 depending on sample and variable compositions. Above the differences in variable selection (Ohlson model nine variables which are different from the five Altman variables) produce unfair comparisons in predictive ability of the models. Thus, I am going to create testing samples of 20 companies each in order to

i) analyse the classification patterns by reducing the differences in variable selection, Section 7.1;
ii) nullify the differences in the different bankrupt to non-bankrupt mixing ratios (Altman estimation sample used a 1:1 ratio whereas the Ohlson estimation sample used a 1:20 ratio) in the estimation samples of the bankruptcy models, Section 7.2;
iii) minimize the effect of the industry heterogeneity, Section 7.3; and
iv) test the firm size dependence on classification patterns, Section 7.4.

7.1 Predictive ability of MDA and Logit models when both models use Altman's five variables

Altman (1968) used five variables which closely related to total assets while Ohlson (1980) used nine variables including liabilities-related financial ratios and a parameter to measure changes in net income. In general, the selection of variables in bankruptcy prediction study is based on pragmatic (statistical or availability-driven), not theoretical, considerations. However, the choice of variables which go into a model will eventually determine the predictive power of that model. In order to ensure comparable results between the prediction power of two models, the models thus should be based on the same variables.

Sample firms and variables

For a training sample, the same bankrupt firms and non-bankrupt firms which were selected to re-estimate the MDA and Logit models (5.4.4) were selected. Thus, the variables for the Logit model are changed to Altman's five variables, WC, RE, EBIT, ME, and SA, while there are non changes to the estimation of the MDA model. Recall that the proportion of bankrupt and non-bankrupt firms within the training samples is 1:1 and 1:20 for the MDA and Logit models, respectively.

In order to ensure comparability of predictive ability, the same sample was applied to both models. The results for the MDA model in are given in 6.1.1. The details for the re-estimation of the logit model are given below.

Re-estimated Logit model

Group means

	WC	RE	EBIT	SA	logME
0	0.02455292	0.08569277	0.14944898	1.278779	-1.180852
1	-0.05879417	-0.25349568	0.03409139	1.484594	-2.152679

1 = bankrupt firms, 0 = non-bankrupt firms

Model fitting

```
Deviance Residuals:
   Min       1Q     Median      3Q      Max
-1.4133  -0.3343   -0.2687   -0.2114   2.9032

Coefficients:
              Estimate  Std. Error   z value   Pr(>|z|)
(Intercept)   -3.75241  0.21312      -17.607   < 2e-16 ***
WC            -0.13988  0.42534      -0.329    0.742256
RE            -0.03915  0.13057      -0.300    0.764293
EBIT          -1.63110  0.57370      -2.843    0.004467 **
logME         -0.31460  0.05754      -5.467    4.57e-08 ***
SA             0.28852  0.08489       3.399    0.000677 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 804.07  on 2099  degrees of freedom
Residual deviance: 747.48  on 2094  degrees of freedom
AIC: 759.48

Number of Fisher Scoring iterations: 6
```

Evaluation of predictive ability

Matrix 7-1 and 7-2 show that the predictive accuracy of the two models is the same at 60%. However, the Type 1 error rate is 10% lower for the logit model than it is for the MDA model. On the other hand, the Type 2 error rate of the MDA model is 70%, which is 10% lower than that of the Logit model.

Matrix 7-1 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	6	4
Prediction accuracy		60%	
Type 1 error		20%	
Type 2 error		60%	

Matrix 7-2 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	9	1
	BKR	7	3
Prediction accuracy		60%	
Type 1 error (FP)		10%	
Type 2 error (FN)		70%	

For the error pattern of two models, Matrix 7-3 and 7-4 show that the MDA model made six Type 1 classification errors on sample firms 2, 3, 5, 6, 8 and 9 while the Logit model made seven Type 1 classification error on the same sample firms 2, 3 and 5-9. For the Type 2 classification error, both models made classificaion error on sample firm 15 but the MDA model made one more classification error on sample firm 18. Given these misclassification patterns, it is not possible to judge which model is better.

Matrix 7-3 Error patterns of MDA model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	0	1	0	0	1	0	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	TP	FP	FP	TP	FP	FP	TP	FP	FP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	0	0	1	0	0	1	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	TN	FN	TN	TN	FN	TN	TN

Matrix 7-4 Error Pattern of Logit model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	0	1	0	0	0	0	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	TP	FP	FP	TP	FP	FP	FP	FP	FP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	0	0	1	0	0	0	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN

7.2 Matching the sample ratio of bankrupt firms vs non-bankrupt firms

The original training dataset which Altman (1968) used consisted of 33 bankrupt and 33 non-bankrupt firms. Evidently, the proportion of bankrupt and non-bankrupt firms is for the

MDA model 1:1. Begley et al. (1996) followed the same proportion rule, and selected 100 bankrupt firms and 100 non-bankrupt firms to create the training dataset. On the other hand, the training dataset which Ohlson (1980) created consists of 105 bankrupt firms and 2,058 non-bankrupt firms. The proportion is approximately 1:20. Begley et al. (1996) followed the same proportion, and selected 100 bankrupt firms and 2000 non-bankrupt firms.

A valid consideration is that proportion of bankrupt and non-bankrupt firms in the testing and particularly the testing samples, should mimic business practice. Academic work however is not concerned with firm entry and exit dynamics and chooses these arbitrary mixing ratios. Here I at least control for the same ratio to test the MDA and Logit models.

1) Sample firms and variables

The proportion of bankrupt and non-bankrupt for both training and testing sample dataset is assumed to be 1:1. The same 100 bankrupt firms used to estimate the models are being used and randomly matched with 100 non-bankrupt firms. For the testing sample, 10 bankrupt firms from the originally 64 bankrupt firms are selected and matched with 10 non-bankrupt firms. The matching procedures follow the rules given in Chapter 5.3. The re-estimated Logit model parameter estimates are given below

Model fitting

```
Deviance  Residuals:
     Min          1Q      Median          3Q          Max
-1.95337   -0.84941   -0.05443    0.86577    2.05139

Coefficients:
                   Estimate   Std.  Error          z  value    Pr(>|z|)
(Intercept)        -0.45310   0.39834              -1.137      0.255346
WC                 -0.80375   1.05760              -0.760      0.447266
RE                 -1.21080   0.55747              -2.172      0.029858  *
EBIT               -2.80475   1.56675              -1.790      0.073427  .
SA                 0.09918    0.17137              0.579       0.562762
logME              -0.53394   0.15286              -3.493      0.000477  ***
---
Signif.  codes:   0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ' '  1

(Dispersion  parameter  for  binomial  family  taken  to  be  1)

    Null  deviance:  277.26   on  199   degrees  of  freedom
Residual  deviance:  206.45   on  194   degrees  of  freedom
AIC:  218.45

Number  of  Fisher  Scoring  iterations:  5
```

Evaluation of predictive ability

Matrices 7-5 and 7-6 give the results of the predictive ability of the MDA and Logit models. The Logit model has a better accuracy rate of 75% while the accuracy rate MDA model is 60%. When comparing the Type 1 error rate, MDA and Logit model shows same result at 20% while the Type 2 error of Logit model is half of the MDA model (30%). According to these results, it can be concluded that the Logit model has better predictive ability.

Matrix 7-5 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	6	4
Prediction accuracy		60%	
Type 1 error		20%	
Type 2 error		60%	

Matrix 7-6 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	3	7
Prediction accuracy		75%	
Type 1 error (FP)		20%	
Type 2 error (FN)		30%	

Considering the detailed misclassification patterns displayed in Matrix 7-7 and 7-8, above conclusion can be re-inforced. The Logit model commits the same classification mistakes as the MDA model does. However, the MDA model additionally misclassifies sample firms 2, 5 and 6.

Matrix 7-7 Error pattern of MDA model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	0	1	0	0	1	0	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Accuracy	TP	FP	FP	TP	FP	FP	TP	FP	FP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	0	0	1	0	0	1	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	TN	FN	TN	TN	FN	TN	TN

Matrix 7-8 Error pattern of Logit model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	1	0	1	1	1	1	0	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Accuracy	TP	TP	FP	TP	TP	TP	TP	FP	FP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
MDA	0	0	0	0	1	0	0	1	0	0
Logit	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	TN	FN	TN	TN	FN	TN	TN

7.3 The effect of the industry SIC code

Begley et al. (1996) restricted the industry SIC code which applied to sample bankrupt firms to be less than 4000 or between 5000 and 5999. The reason for this was to mimic the restriction imposed on drawing observations from one industry and create similar sample datasets such as in Altman (1968) and Ohlson (1980). Altman focused on manufacturing firms and Ohlson focused on industrial firms.

Table 7-2 shows the number of bankrupt firms by SIC code and within the 1980 to 1989 analysis window, and Table 7-3 a more detailed yearly breakdown. The biggest division is the manufacturing sector which contains 157 bankrupt firms throughout the period of analysis. The second largest sector is retail, followed by service sector. The replicated sample (5.4.3) included the largest sector, but also include A (agriculture, Forestry, and Fishing), B (Mining), C (Construction), F (Wholesale Trade), and G (Retail Trade). The problem of creating the replicated sample is that six out of ten divisions are compressed in one sample. To control for specific rules (reporting standards, industry regulation) that may be adopted for one sector but not for companies outside a given sector, I focus here on the manufacturing industry only.

Table 7-1 - Number of bankrupt firms by SIC code (1980-1989)

Divison	SIC code	Number of Firms
A: Agriculture, Forestry, And Fishing	SIC < 1000	2
B: Mining	1000<SIC<1499	37
C: Construction	1500<SIC<1800	3
D: Manufacturing	2000<SIC<4000	157
E: Transportation, Communications, Electric, Gas, And Sanitary Services	4000<SIC<5000	21
F: Wholesale Trade	5000<SIC<5200	20
G: Reatil Trade	5200<SIC<6000	47
H: Finance, Insurance, And Real Estate	6000<SIC<7000	11
I: Services	7000<SIC<9000	39
J: Public Administration	9000<SIC<10000	4

Table 7-2 - Number of bankrupt firms by SIC code

	Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Division	A: Agriculture, Forestry, And Fishing	0	0	0	0	0	0	2	0	0	2
	B: Mining	2	3	3	6	6	5	5	6	1	37
	C: Construction	1	0	0	0	0	2	0	0	0	3
	D: Manufacturing	8	11	9	10	9	20	19	64	7	157
	E: Transportation, Communications, Electric, Gas, And Sanitary Services	3	1	1	3	1	0	2	10	0	21
	F: Wholesale Trade	1	1	0	3	2	1	0	12	0	20
	G: Reatil Trade	1	2	5	3	8	2	5	20	1	47
	H: Finance, Insurance, And Real Estate	1	0	0	2	1	2	3	1	1	11
	I: Services	0	2	0	3	5	11	2	14	2	39
	J: Public Administration	0	0	0	0	0	1	0	3	0	4
	Total	17	20	18	30	32	44	38	130	12	341

After implementing the total asset rule (larger the \$10Million and smaller than \$200Million), the total number of bankrupt manufacturing firms that enter my estimation and testing dataset are as follows:

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of BKR firms	7	10	8	10	8	19	16	62	7	147

From above sample, 100 bankrupt firms are selected to create 50% of the training dataset.

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of BKR firms	5	7	5	7	5	13	11	42	5	100

For matching (by size) the bankrupt firms, the total number of non-bankrupt firms are as follows,

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of NBKR firms	189	191	192	188	194	196	197	185	18	1550

from which I then create a randomly drawn paired sample.

Thus, I have a 200 firm training dataset and I use Altman's five variables WC, RE, EBIT, log(ME), and SA, for both MDA and Logit models. For creating the testing sample, the same criteria apply and I choose to continue with the small 20 firm training dataset.

The details of the re-estimation of the MDA and Logit models follows:

Group means

	WC	RE	EBIT	SA	logME
0	-0.0093	-0.1800	0.0580	1.131	-0.6382685
1	-0.0766	-0.5551	-0.0344	1.491	-2.1526787

Model fitting (MDA – Coefficients)

```
Coefficients of linear discriminants:
                 LD1
WC               0.44344228
RE               0.01235486
EBIT             -2.06886034
logME            -0.38362758
SA               0.76180041
```

Model fitting (Logit)

```
Call:
glm(formula = BKR ~ . - ME, family = "binomial", data = TRdata)

Deviance Residuals:
    Min       1Q    Median      3Q      Max
-2.0704  -0.8894  -0.0458   0.8574   2.2768

Coefficients:
            Estimate   Std. Error   z value   Pr(>|z|)
(Intercept)  -2.0538    0.4144      -4.956    7.20e-07 ***
WC           -0.3126    0.5878      -0.532    0.59484
RE           -0.1744    0.2038      -0.856    0.39208
EBIT         -0.5598    0.9596      -0.583    0.55965
SA            0.7283    0.2500       2.913    0.00358 **
logME        -0.7956    0.1393      -5.711    1.12e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 277.26  on 199  degrees of freedom
Residual deviance: 211.62  on 194  degrees of freedom
AIC: 223.62
```

Evaluation of predictive ability

Matrices 7-9 and 7-10 evaluate at the aggregate level the results of the predictive ability of MDA and Logit models when both models are estimated from a training dataset which consists of bankrupt and non-bankrupt firms from only Division D (manufacturing sector) under SIC classification. In this instance, the Logit model has a slightly better accuracy rate of 85% than the MDA model (80%). However, at the Type 1 and Type 2 error level, the classification accuracy is mixed.

Matrix 7-9 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	9	1
	BKR	3	7
Prediction accuracy		80%	
Type 1 error		10%	
Type 2 error		30%	

Matrix 7-10 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	1	9
Prediction accuracy		85%	
Type 1 error (FP)		20%	
Type 2 error (FN)		10%	

The detailed error patterns shown in Matrices 7-11 and 7-12 show that both the MDA and Logit models made a Type 1 classification error on sample firm 4, and a Type 2 classification error on sample firm 14. The MDA model made two further Type 1 classifficaion errors while the Logit model made one more Type 2 classificationerror. Based on this instance of a training sample, it is not possible to judge which of the two models may have better overall predictive accuracy now and even less so should a new training sample be applied. In situations where either Type 1 or Type 2 errors are deemed to be much more costly, one could, based on the above 20 firm training set, determine which model be preferable.

Matrix 7-11 Error patterns of MDA model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	0	1	1	0	1	1	0	1	1	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	FP	TP	TP	FP	TP	TP	FP	TP	TP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	0	1	0	0	0	0	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	TN	FN	TN	TN	TN	TN	TN	TN

Matrix 7-1 Error patterns of Logit model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	1	1	0	1	1	1	1	1	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	TP	TP	TP	FP	TP	TP	TP	TP	TP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	1	0	0	1	0	0	0	0	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	FN	TN	TN	FN	TN	TN	TN	TN	TN	TN

7.4 Effect of firm size (total asset) on classification patterns

Begley et al. (1996) included in their sample firms with total assets exceeding $10Million. From my dataset, the reason may be due to quite a few bankruptcy filing dates for smaller firms are missing. Any given economy is dominated by small and medium sized firms. It is therefore useful to testif the predictive performance of the MDA and Logit model on the largest firms of the US economy also apply to (somewhat) smaller firms.

Table 7-4 shows the size distribution of bankrupt (BKR) and non-bankrupt (NBKR) firms.

Table 7-4 Categorizing Compustat firms according to size during 1980-1988. BKR firm total asset is determined from the last financial statement information (as per Chapter 5). NBKR data displays the number of firm year observations.

	BKR		NBKR	
Total Assets (mil)	Number of Firms	%	Number of Firms	%
TA<10	111	33%	762	18%
10<TA<50	128	38%	1116	27%
50<TA<100	45	13%	476	11%
100<TA<200	28	8%	483	12%
TA>200	29	9%	1303	31%
Total	341	100%	4140	100%

To re-estimate the MDA and Logit models, I adopt now the most comparable samples and models: Both models contain the five original Altman variables, and both samples have a 1:1 mixing ratio between bankrupt and non-bankrupt firms, all of which have total assets less than $10Million.

The total number of bankrupt firms which match the criteria above are shown here:

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of BKR firms	3	4	6	7	15	21	12	40	3	111

I then choose randomly 100 bankrupt firms to create a training sample:

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of BKR firms	2	3	5	6	13	19	11	38	3	100

And prepare from the non-bankrupt firms

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	Total
Number of NBKR firms	59	69	95	93	105	104	114	109	14	762

a paired sample where the same number of non-bankrupt firms are randomly drawn to also match the appropriate year.

The details of the re-estimation of the MDA and Logit models follows:

Group means

```
Group  means:
        WC              RE          EBIT        SA          logME
0       -0.0649         -0.9138     -0.02308    1.035       -0.6382
1       -0.0364         -1.2954     -0.14668    1.152       -2.1526
```

Model fitting (MDA – Coefficients)

```
Coefficients  of  linear  discriminants:
                 LD1
WC         0.06246958
RE        -0.08350852
EBIT      -0.56242560
logME     -0.67900119
SA         0.25382690
```

Model fitting (Logit)

```
Call:
glm(formula  =  BKR  ~  .  -  ME,  family  =  "binomial",  data  =  TRdata)

Deviance  Residuals:
     Min           1Q       Median          3Q          Max
-2.13758   -0.94743   -0.06102      0.87651     2.23768

Coefficients:
                      Estimate    Std.  Error           z  value     Pr(>|z|)
(Intercept)           -1.58369    0.36218               -4.373       1.23e-05  ***
WC                    0.05789     0.16888               0.343        0.7318
RE                    -0.10676    0.13384               -0.798       0.4251
```

```
EBIT                 -0.61208   0.44282            -1.382    0.1669
SA                    0.28673   0.16797             1.707    0.0878  .
logME                -0.80429   0.13460            -5.976    2.29e-09  ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 277.26  on 199  degrees of freedom
Residual deviance: 218.72  on 194  degrees of freedom
AIC: 230.72

Number of Fisher Scoring iterations: 4
```

Evaluation of predicting ability

Matrices 7-14 and 7-15 show the predictive accuracy of the two models. The Logit model has a better rate at 85% than the MDA model at 75%. At the individual error level, both models show the same Type 1 error rate. However, with respect to the Type 2 error, the Logit model performs better with a 10% error rate than the 30% error rate for the MDA model. At this level of analysis, it can be concluded that the Logit model has a better predictive ability than MDA model.

Matrix 7-14 Confusion matrix of MDA model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	3	7
Prediction accuracy		75%	
Type 1 error		20%	
Type 2 error		30%	

Matrix 7-15 Confusion matrix of Logit model

		Predict	
		NBKR	BKR
Actual	NBKR	8	2
	BKR	1	9
Prediction accuracy		85%	
Type 1 error (FP)		20%	
Type 2 error (FN)		10%	

Now considering at the detailed level of individual firm classifications, Matrices 7-16 and 7-17 show that the Type 1 error comes from miss-classifying the same two companies (sample firms 13 and 15). However, with respect to the Type 2 error, both MDA and Logit models made classification errors on different sample firms. From such results, we cannot conclude superiority of one method over another due to the apparent data-dependency of results.

Matrix 7-2 Error pattern of the MDA model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	1	1	0	1	1	1	0	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	FP	TP	TP	FP	TP	TP	FP	TP	TP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	1	0	1	0	0	0	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	FN	TN	FN	TN	TN	TN	TN	TN

Matrix 7-3 Error pattern of the Logit model

BKR	1	2	3	4	5	6	7	8	9	10
Predict	1	0	1	1	1	1	1	1	1	1
Actual	1	1	1	1	1	1	1	1	1	1
Class	TP	FP	TP	TP	TP	TP	TP	TP	TP	TP
NBKR	11	12	13	14	15	16	17	18	19	20
Predict	0	0	1	0	1	0	0	0	0	0
Actual	0	0	0	0	0	0	0	0	0	0
Class	TN	TN	FN	TN	FN	TN	TN	TN	TN	TN

Chapter 8. Summary and conclusions

8.1 Empirical result

In this research, I examined the results of Begley et al. (1996) who concluded that the Logit model is superior over the MDA model in bankruptcy prediction. I replicated samples based on the criteria Begley et al. employed. I re-estimated the MDA and Logit models using a training data set and I assessed a predictive ability of the two models using various testing samples. When the Type 1 and Type 2 error rates are compared at the aggregate level in the confusion matrix, the results were often, but not always, consistent with the study of Begley et al. When comparing classification error patterns in matrix format however, I provide evidence that conclusions at the aggregate Type 1 and Type 2 error-level may be misleading. I have shown the the better predictive ability to correctly classify firms in a testing sample to two categories of bankrupt and non-bankrupt firms depends on sample composition. Thus, any judgment about model superiority would be constrained by i) the choice of the training data set, ii) the choice of independent variables that go into a model, and iii) the choices made around creating the testing samples. For example, I have found no evidence why the selected bankrupt vs non-bankrupt mixing ratios are appropriate and how such drive the predictive performance of bankruptcy models. Should these ratios mimic the entry and exit dynamics of firms in the economy?

In order to answer my research question 'Do the reported results of the MDA and Logit models in the literature with respect to bankruptcy prediction hold when firm-specific classification patterns are considered?', I created two samples using the same Altman's five variables under different proportion of bankrupt and non-bankrupt of 1:1 and 1:20 for MDA and Logit model, repectively. I re-estimated the Logit models and compared a predictive ability between the two models using two testing samples with 256 and 20 matched 1:1 bankrupt vs non-bankrupt firms. The results are inconclusive. Thus, with respect to the research question I have demonstrated, by providing a negative result, that classification patterns at the individual firm level will question conclusions one might draw at the usually applied aggregate Type 1 and Type 2 level.

I then have created further samples to generate a more just context for the comparison of the predictive ability of the MDA and Logit models. Clearly, when comparing any performance, such as for example in sports, one is to agree on as similar (or fair) conditions of departure such that only intrinsic performance is being assessed. E.g., nobody would be interested in formula one races where the same car chassis is provided but one team was allowed to mount a V12 engine running on some high ignition fuel, the other team a solar driven battery propulsion system. Therefore I have used the same five original Altman variables (working capital/total assets; retained earnings/total assets; earnings before interest & taxes/total assets; sales/total assets; and market value of equity/book value) in both models and I have used the same proportion (1:1) of bankrupt and non-bankrupt firms to estimate the model parameters. Under such conditions my results show that particularly under the microscopic lense of individual classification patterns, no clear favourite in terms of predictive ability emerges.

Then in a series of further restrictions (firm size and industry) on the estimation and testing sample composition, the previous inconclusive results with respect to superior predictive ability have been obtained.

8.2 Further study

The ideas presented in this study can be explored in a more robust setting. For example, I have limited my data to the 1990ies, US publicly traded companies, within particular industries and particular sizes. I also have not tested other than the original Altman and Ohlson model parameters. Future work would clearly also test for all the other claims levelled in the bankuptcy literature of which model performs better – in particular when classical statistical models of the 70ies and 80ies are being compared with computationally advanced, modern algorithms.

The development of technology also enables us to implement Monte Carlo ideas so that hundreds of combinations of bankrupt and non-bankrupt firm estimation and testing samples are created. This would elevate the judgment based on particular instances of an estimation and testing sample to a more relative, better assessment of model performance.

Lastly, this research has not really incorporated the model fitting characteristics, i.e., the statistical assumptions of the model have been of no concern. Yet clearly, one would not expect, say, the MDA model to outperform the Logit model if the statistical specification tests flagged concerns in the former and none were raising concerns for the latter.

Chapter 9. References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The* journal *of finance*, *23*(4), 589-609.

Altman, E. I. & Eisenbeis, R. A. (1978). Financial applications of discriminant analysis: a clarification. Journal of Financial and Quantitative Analysis, 13(1), 185-195.

Altman, E. I. (1980). Commercial bank lending: Process, credit scoring, and costs of errors in lending. Journal of Financial and Quantitative Analysis, 15(4), 813-832.

Altman, E. I. (1983). Corporate financial distress: A complete guide to predicting, avoiding, and dealing with bankruptcy. Wiley.

Altman, E. I. (2000). Predicting financial distress of companies: revisiting the Z-score and ZETA models. *Stern School of Business, New York University*, 9-12.

Altman, E. I., & Hotchkiss, E. (2010). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (Vol. 289): John Wiley & Sons.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2014). Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2016). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting*.

Aly, I. M., Barlow, H., & Jones, R. W. (1992). The usefulness of SFAS No. 82 (current cost) information in discriminating business failure: an empirical study. *Journal of Accounting, Auditing & Finance, 7*(2), 217-229.

Back, B., Laitinen, T., Sere, K., & van Wezel, M. (1996). Choosing bankruptcy predictors using discriminant analysis, logit analysis, and genetic algorithms. *Turku Centre for Computer Science Technical Report, 40*, 1-18.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting*

Review, 38(1), 63-93.

Baldwin, J., & Glezen, G. W. (1992). Bankruptcy prediction using quarterly financial statement data. *Journal of Accounting, Auditing & Finance, 7*(3), 269-285.

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

Beaver, W. H. (1968). Alternative accounting measures as predictors of failure. *The Accounting Review, 43*(1), 113-122.

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of accounting studies, 1*(4), 267-284.

Douglas, E., Lont, D., & Scott, T. (2014). Finance company failure in New Zealand during 2006–2009: Predictable failures? *Journal of Contemporary Accounting & Economics, 10*(3), 277-295.

Falta, M., & Willett, R. (2013) The Impact of Exponential Growth Variables on Regression Models of the Market Accounting Relation. *Journal of Accounting Auditing and Finance,* 28(3), 243-272.

Eisenbeis, R. A. (1977). Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. *The journal of finance, 32*(3), 875-900. doi:10.1111/j.1540-6261.1977.tb01995.x

Grice, J. S., & Ingram, R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research*, 54(1), 53-61.

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of accounting studies, 9*(1), 5-34.

Ho, C.-Y., McCarthy, P., Yang, Y., & Ye, X. (2013). Bankruptcy in the pulp and paper industry: market's reaction and prediction. *Empirical Economics, 45*(3), 1205-1232.

Iturriaga, F. J. L., & Sanz, I. P. (2015). Bankruptcy visualization and prediction using neural networks: A study of US commercial banks. *Expert Systems with Applications, 42*(6), 2857-2869.

Kim, M.-J., & Kang, D.-K. (2012). Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction. *Expert Systems with Applications*, 39(10), 9308-9314.

Luoma, M., & Laitinen, E. (1991). Survival analysis as a tool for company failure prediction. *Omega, 19*(6), 673-678.

Neophytou, E., & Molinero, C. M. (2004). Predicting corporate failure in the UK: a multidimensional scaling approach. *Journal of Business Finance & Accounting, 31*(5-6), 677-710.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research, 18*(1), 109.

Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6(1), 34-45.

Wu, D., Liang, L., & Yang, Z. (2008). Analyzing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis. *Socio-Economic Planning Sciences*, 42(3), 206-220.

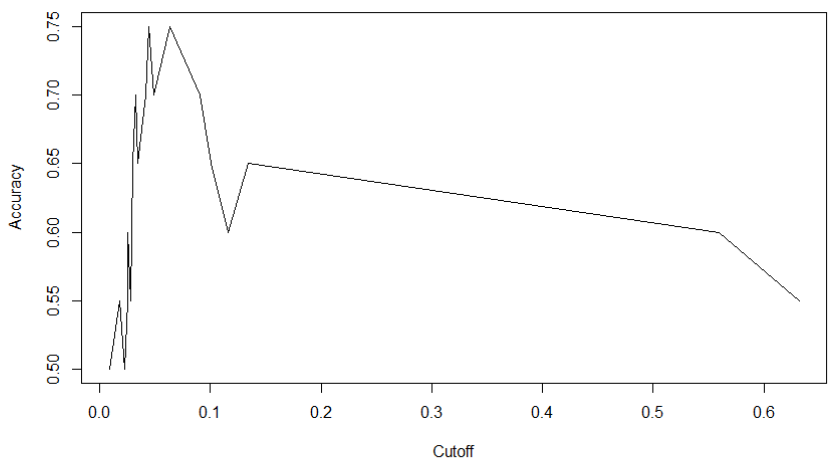
Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 59-82.

Chapter 10. Appendix 1 – Cut-off points and ROC curves

Below, I explain the procedure of how I determined the optimal cut-off points for the various logit models that I have estimated in Chapters 6 and 7. A more detailed description of cut-off determination in R with the rocr() package is given in https://cran.r-project.org/web/packages/ROCR/ROCR.pdf. Here I use two evaluation graphs which in tandem yield the numerical values for the various cut-offs .

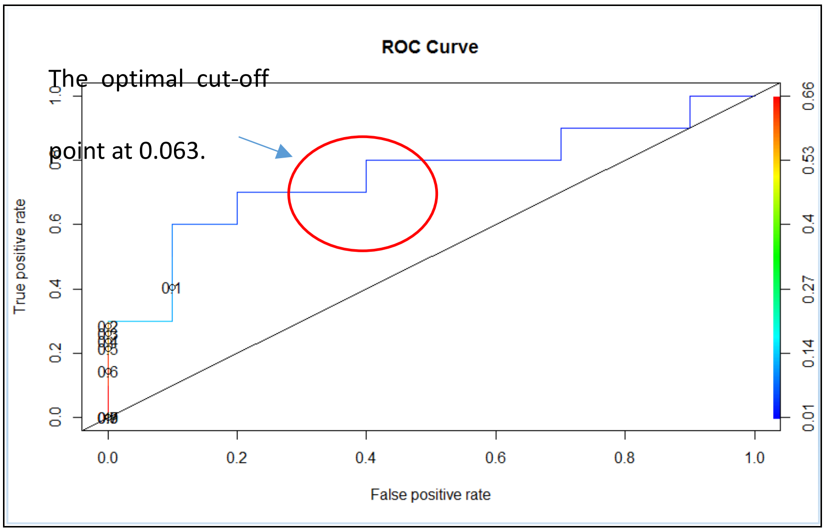
1. Cut-off point and ROC curve for Logit model to replicate Begley et al. (1996)

The graph below illustrates can be obtained via the command performance(ROCRPred, x) which displays the relationship btween the predictive accuracy (Accuracy) and cut-off-points (Cutoff). The particular instance below applies to the estimated Logit model in Chapter 5.1.2. The optimal cut-off point can be found at the highest point of the Accuracy level (75%), which is 0.0633.

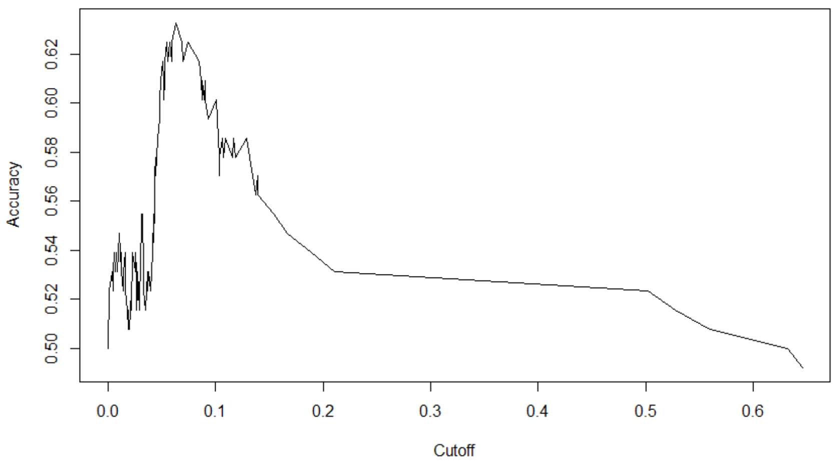


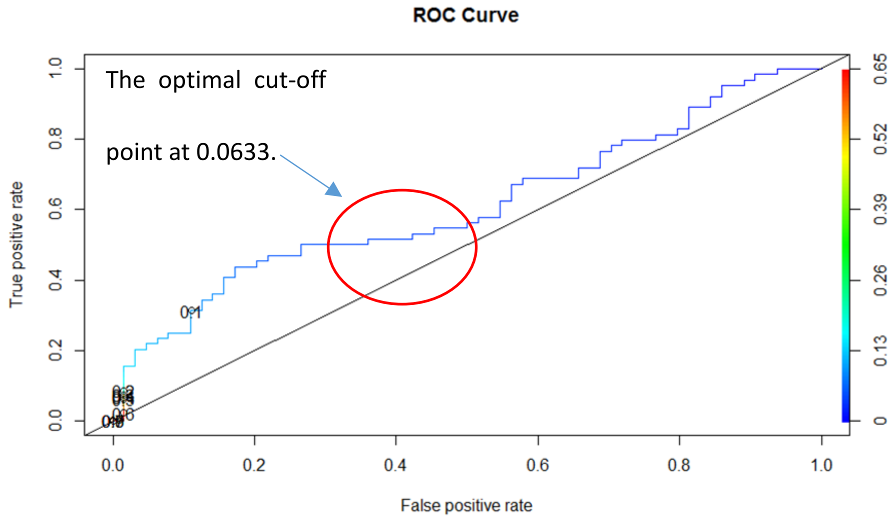
The following graph shows the ROC curve. ROC curves show the trade-off between the true positive rate and false positive rate of a given Logit model at different cut-off points. The top left corner of the plot is the ideal point for the model in classifying firms into bankrupt and non-bankrupt firms. The optimal cut-off point always lies between 0 and 1, and is obtained at the closest point that ROC curve follows to the top left corner which is the point where Type 1 and Type

2 error rates are jointly minimized. As ROC curve goes left corner, cut-off points increase. The following graph shows the cut-off point at 0.0633, which can be found at the top left of the ROC curve. I have circled (red) the range of optimal cut-offs.

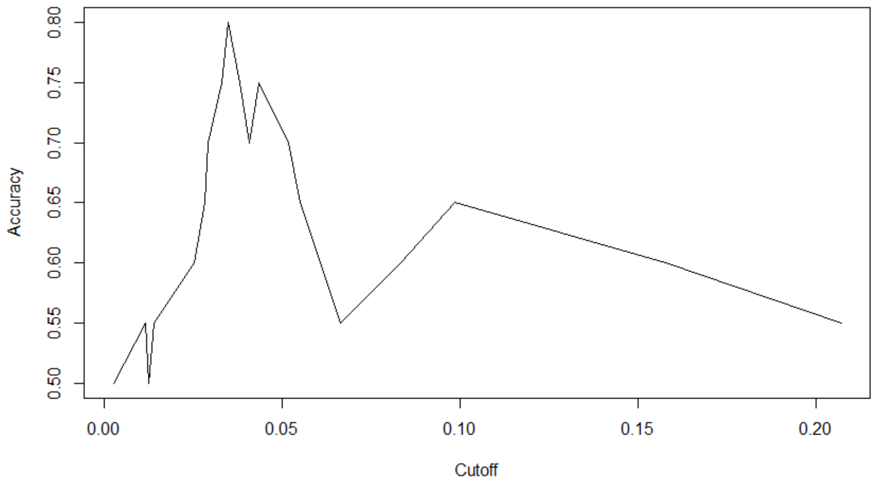


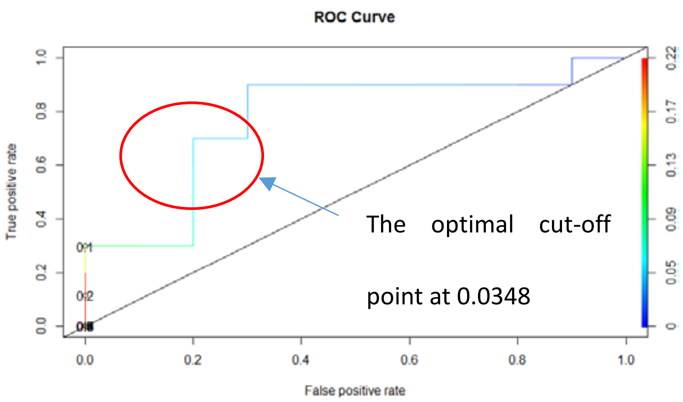
2. Determining the cut-off point for the replication sample of the Begley et al. (1996) study when the number of firms in the testing sample is 128 (64 bankrupt firms and 64 non-bankrupt firms). The optimal cut-off point is 0.0633 at 63% of predictive accuracy.



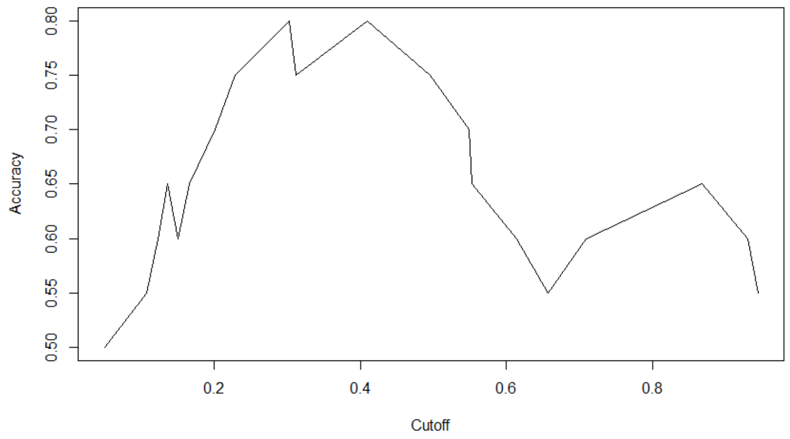


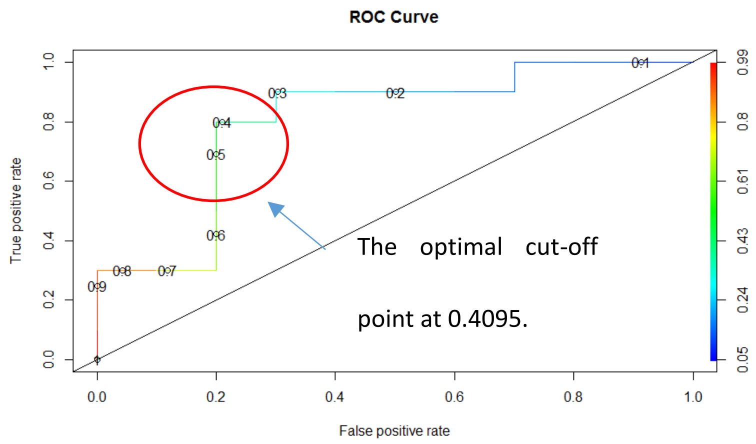
3. Determining the cut-off point and ROC curve when two models are estimated by the same set of independent variables (Chapter 7.1). The optimal cut-off point is 0.0348 at 85% of predictive accuracy.



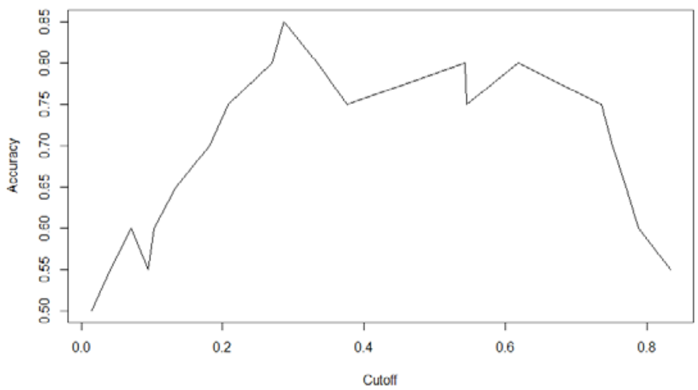


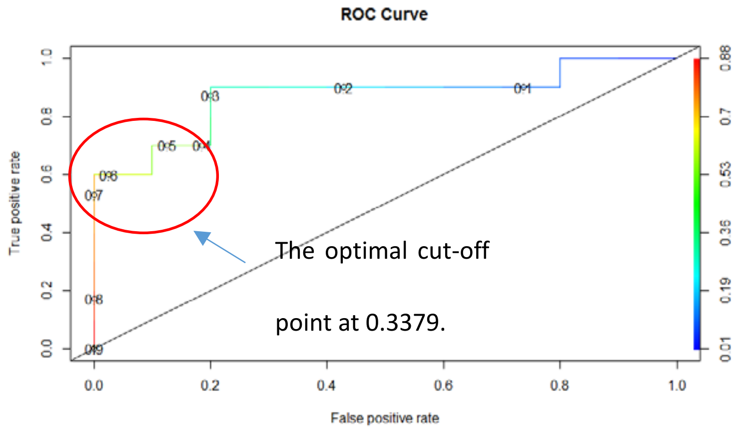
4. Determining the optimal cut-off point when two models are estimated by the same proportion of bankrupt and non-bankrupt firms (Chapter 7.2). The optimal cut-off point is 0.4095 at 80 % of predicitive accuracy.





5. Determining the cut-off point when two models are estimated by sample firms only from manufacturing sector (Chapter 7.3). The optimal cut-off point is 0.3779 at 70% of predictive accuracy.





6. Determining the cut-off point when two models are estimated by sample frims whose total assets are less than $10Million (Chapter 7.4). The optimal cut-off point is 0.4889 at 90% of predictive accuracy.