SCHOOL OF ECONOMICS
AND MANAGEMENT
Lund University

Master Thesis
Spring 2007

MANAGING CREDIT RISK:

ASSESSING THE PROBABILITY OF CORPORATE BANKRUPTCY USING QUANTITATIVE RISK ANALYSIS

Supervisor
Göran Anderson

Authors
David Granholm
Theodoros Goumas

ABSTRACT

Title: Managing Credit Risk: *Assessing the Probability of Corporate Bankruptcy using Quantitative Risk Analysis.*

Seminar date: 2007-06-04

Course: BUS860**,** Master Thesis in Business Administration, 10 Swedish credits (15 ECTS)

Authors: David Granholm and Theodoros Goumas

Advisors: Göran Anderson

Key words: Credit risk, bankruptcy, bankruptcy prediction, financial ratios, discriminant analysis, logistic regression, neural networks.

Purpose: The main purpose of this master thesis is to assess the probability of bankruptcy for firms in the Swedish market. Bankruptcy prediction can be valuable tool for managing credit risks efficiently. Further, the thesis evaluates which of the employed quantitative models that are best suited for bankruptcy prediction.

Methodology: Gathering financial ratios for 113 defaulted firms and equally many non-defaulted firms in Sweden between 2004-2007, two groups are created; An in-sample model specification group and a holdout-sample group for model testing. Three quantitative methods are used to assess the probability of bankruptcy: (i) Discriminant Analysis; (ii) Logistic Regression and; (iii) Neural Networks.

Theoretical perspective: The theoretical framework largely involves prior research in the field of bankruptcy prediction. Further, firm credit ratings and Swedish bankruptcy law are studied.

Results: We can with the models at hand predict bankruptcy in the range of 74% to 87%. The predictive accuracy of this study is in line with prior research.

Conclusions: Quantitative risk analysis can benefit a corporate creditor in assessing the probability of a loan applicant, thus the models provided can be a useful tool in managing credit risk more efficiently. We conclude that standard logistic regression should be applied in predicting corporate bankruptcy.

TABLE OF CONTENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In this initial section the reader is introduced to the subject that is assessed in this thesis, starting with a background description of the problem. Further specification are given of the particular problem that is examined and what limitations that are set up in the study.

1.1 BACKGROUND

Managing credit risks might be the single largest field of business for any commercial bank. The decision whether to grant credit or not to a client might make the difference between conducting business in a efficient manor for the banking industry; On the one hand it wouldn't be beneficial for the bank to grant credit to a corporate client if there is a high possibility that continuous liability payments can not be met and on the other hand it would neither be good business conduct to neglect a client with high possibility of fulfilling the interest payments.

Financial distress enters into a company when it has a problem with fulfilling its continuous liability payments. If the firm does not solve the distress situation it could very easily enter a bankruptcy state, where the firm's assets would be liquidated and divided among its liability holders. This could potentially mean that the holders of the liability loose a large amount of money due to generally two factors; i) The process of liquidating and obtaining lost capital is time consuming and costly and; ii) The risk of the creditor not obtaining the full capital at stake while the firm was operating in the going-concern state. This is an unwanted event in the eyes of the credit grantor, which stands to lose large sums. Herein lays the incentive for creditors and for that matter other stakeholders, to carefully assess its client's probability of defaulting in order to minimize their credit risk exposure. A good estimator of what characterizes a firm that has a high probability of defaulting could be a helpful tool in managing credit risks in an efficient manor.

Assessing the probability of bankruptcy has been a widely explored field in the research community. When the era of industrialism started to take off, and more and more firms were forced to borrow money for their required capital investments, the banking business flourished. This "boom" in the business also called for more accurate methods to asses the risk exposure that was imposed on the lenders. Starting off with in-house expertise, that is calculating the probability of default using prior qualitative knowledge, the methods have in recent years been significantly refined. The starting point assessing corporate bankruptcy using quantitative data is an article from 1966 by William H Beaver, who uses ratio analysis between healthy and distressed firms to approximate the probability of bankruptcy. This approach has later on been discarded as far too simplified but there are researchers that find it to be a useful tool for bankruptcy prediction (see Arshad 1985). However, these simplified models have little predictive ability and applying these models should be precautionary, thus more refined quantitative models are currently applied.

Probably the most famous article on the topic was presented by Edward Altman in 1968, who applies a multiple discriminant analysis on a set of financial ratios between two sample groups. The number given by the discriminant function, the credit score, is known as the Z-score. Credit scoring type models specifies a number of financial and performance ratios for the firm which impacts the probability of bankruptcy. The final parsimonious model is comprised of the ratios that contribute the most to the bankruptcy. A common observation is that the leverage is a significant determinant of the firm survivorship (Altman 1968; Johnsen et al 1994; Hillegeist et al 2004; Dietrich et al 2005; Hua et al 2007), i.e. PD=f(D/E)[1], which has also traditionally been intuitively used as a proxy for the probability of financial distress.

Another way, apart from calculating credit scores, has been the application of the Black-Scholes-Merton option pricing model (BSM) to estimate the probability of bankruptcy. Option pricing models, which are frequently used by large credit firms, also yield a number to which an interpretation is linked. By analyzing thousands of firms a database is created to rank firms by their creditworthiness and of which AAA is the most creditworthy grade (Kim et al 2004). A major drawback of these types of option pricing models, eventhough proven predicatively robust (Hillegeist et al 1994), is the fact that the BSM requires the firm's equity to be publicly priced. This requirement puts a major constraint on the types of firms that can be evaluated according to this model, because of the fact that the vast majority of all firm's equities are not traded in the market place. Therefore a wide usage of an option type based model, for example in the banking industry, would be severely crippled by the fact that their day-to-day operations largely or exclusively deals with privately held firms.

The credit scoring type model can however be applied on non-traded firms as well as traded firms. Throughout the years several models have been applied for assessing bankruptcy probabilities. Apart from Multiple Discriminant Analysis (Altman 1968; Lo et al 1986; Holmen 1988; Grice et al 2001; Pompe et al 2005) also Logistic Regression (Ohlson 1980; Mensah 1984; Laitinen et al 2000; Nishikawa 2002) and Neural Networks (Raghupathi 1991; Fletcher et al 1993; Wilson et al 1994; Barniv et al 1997; Yang et al 1999; Zhang et al 1999; Atiya 2001; Baek et al 2003) has frequently been used as bankruptcy prediction models. The three models are by far the most widely recognized and acknowledged bankruptcy prediction models and account for 59,3% of the total model usage for bankruptcy prediction applications prior to 2006 (Aziz et al 2006)[2]. Multiple discriminant analysis alone account for 29,1% of total model application. The models individual superiority has been discussed in the research community but a general consensus seems to be that they are useful tools in assessing the probability of corporate bankruptcy. A study by Back et al (1997) show that neural network were superior to the other two models on 1 and 3 years forecasting, and discriminant analysis was superior on 2 year ahead forecasting. On non of the occasions was the logit model superior to the others. Other studies however show the opposite relation when comparing logit to discriminant analysis that is the logit model has higher predictive classification accuracy than the discriminant analysis (see Kim et al 2006).

All of the three approaches for modelling bankruptcy probability have an overall high predictive accuracy. Aziz et al (2006) reports that prior research concludes that discriminant analysis, logit regression and neural networks have prediction accuracies of 85, 87 and 87 percent respectively. There remains no doubt that the models, if correctly implemented, can be a useful tool in assessing corporate default probabilities and furthermore help a credit grantor more efficiently managing it's credit risk.

[1] PD is the probability of defaulting

[2] In the study by Aziz and Humayon (2006) prior research on bankruptcy prediction from ten selected countries was examined.

1.2 DISCUSSION OF PROBLEM

In the process of assessing firms, according to their creditworthiness, determining the financial soundness of the firm is important. This task is rarely easy because of the, in many cases, complex structure of modern firms. Many researchers claim that calculating the probability of bankruptcy using a firm's financial ratios is a valuable tool (Altman 1968; Ohlson 1980; Zavgren 1985). Particularly assessing default probabilities for a firm conditional on the firm's specific economical and financial characteristics is an important parameter in calculating the expected loss for a certain firm in case of bankruptcy. That is the expected amount of money that a stakeholder stands to lose should the firm be resolved. The high accuracy of quantitative tools in the assessment of bankruptcies makes them interesting in predicting bankruptcies; Studies shows overall predictive accuracies of over 80% using statistic and econometric techniques such as multiple discriminant analysis, logistic regression and neural networks (Altman 1984; Altman 2000; Charalambous et al 2000; Sudheer et al 2004). That is the model is on average accurately categorizing, in 8 cases out of 10, whether the investigated firm belongs to the bankruptcy group or the "healthy" group. The currently popular BSM model to assess corporate bankruptcy suffers from one major drawback, the company's equity has to be publicly traded. This restriction makes it unsuitable as a tool for banks and other financial institutions. We suggest that a credit scoring model should be employed by lenders since these models lack this constraint and it is thus more applicable in bank's day-to-day operations mostly dealing with un-listed firms.

The particular problem of determining bankruptcy probabilities is interesting for the credit grantors such as banks and other institutions. Furthermore, trade credit, is extensively used by firms today for short-term financing of its ongoing operations (Chi et al 2006). This implies that other parties apart from the traditional credit grantor would benefit from information about firm's general financial soundness in the decision-making process of whom to grant trade credit to. The knowledge of which corporate customer that is more likely to not return an outstanding credit is valuable to minimize the supplier's losses. Although, increasingly more firm stakeholders benefit from this knowledge due to more refined techniques in operational financing, the traditional corporate credit grantor such as financial institutes and banks are prevalent and still the main beneficiary of the information of default probabilities.

The formulation of our research problem is two-fold; (I) Is predicting corporate bankruptcy possible using traditional quantitative methods conditional on publicly available firm financial data one year prior to the point of default[3] and; (II) Which of the three approaches: (i) Multiple Discriminant Analysis (ii) Logistic Regression and (iii) Neural Networks is best suited to predict corporate bankruptcy assessing the robustness in predictive accuracy of each model.

[3] Predicting corporate bankruptcy in this sense means develop a quantitative model which on average accurately predicts bankruptcies more than chance.

1.3 PURPOSE

The purpose of this thesis is to provide the reader with a quantitative model to assess the probability of corporate bankruptcy one year ahead of the point of default in the Swedish market. To our knowledge no previous research assessing the probability of corporate bankruptcy applying the three quantitative models stated above has been conducted in the Swedish market.

1.4 LIMITATIONS

The limitations of this study include investigating non-financial Swedish firms which defaulted between the years 2004 to 2007, which is due to the data availability and reliability before this time period and the need to turn to several other sources to obtain the necessary information before these years. Furthermore, financial firms are excluded from the study for two reason: (i) Financial firms have special financial characteristics that significantly differ from other firms which could potentially contribute to errors in the methodological accuracy and; (ii) Certain financial firms, such as large banks, have special government bankruptcy protection so called government bailout which if included leads to the study being biased. There are reports of banks having a probability of bankruptcy of 1×10^{-99} at one point in time during 35 years (Scott 1981).

The study is only conducted on firms which have 30 or more employees, this because of the sometimes blurry and inaccurate financial reporting of smaller firms. Larger firms have a greater possibility to more accurately record the real underlying information of the firms' financials because of the existence of an administrative staff in larger firms. We consider this to be organizations of 30 or more employees.

1.5 Outline of Thesis

Chapter 1 - Introduction

The first chapter introduces the reader to the study and the background for conducting a bankruptcy probability study. Further the specific problem and purpose is specified so that the reader clearly can distinguish what the aim of the study is. A short description of why assessing the probability of corporate default is an interesting and important field of study is also given.

Chapter 2 – Theoretical Framework

This chapter deals of the relevant theories of bankruptcy and bankruptcy prediction. More specifically this section consists of bankruptcy theory, credit risk and other associated theories of relevance. A more thorough overview of prior research is also presented and the main results of the chosen research articles are given. A special focus on the benchmark articles which has largely impacted the bankruptcy research community is highlighted. Furthermore, the three mathematical models that is used in the study is presented and an extensive background of usage and mathematical characteristics is specified. Also, the benefits and drawbacks of each model are presented.

Chapter 3 - Methodology

In this chapter the data collection process is described in detail to enable precise replication, and data of included firms are presented. Also, the examined ratios (variables) are specified with empirical support. Further a thorough specification of the empirical modelling is given.

Chapter 4 – Results

In the fourth chapter the results are presented. The results of all the tested models are carefully discussed and a test for the individual models robustness is conducted with a holdout sample. This validation of the models is used to base which model is best suited for predicting bankruptcies.

Chapter 5 – Analysis

The analysis chapter further deepens the understanding and the interpretation of the obtained results. A general discussion of the model performance and predictability is given and a relevancy analysis is carried out. Further, a comparison to prior research is done and explanations are put forward why we obtain the results at hand.

Chapter 6 – Conclusions

In this section conclusion about our models robustness is presented and the formulation of the problem will be concluded. Also the methodological flaws will be included in this section to give a more accurate overview of the reliability of our results. Further, concluding remarks of the study's applicability is done and our recommendations to the beneficiaries of the model that is presented in this study are specified.

CHAPTER 2

THEORETICAL FRAMEWORK

In the theory section the reader is presented to prior research on the topic including what models that are used and the results that is provided. Further the section contains a theoretical framework of relevant topics in bankruptcy theory such as credit rating and the regulatory framework to understand which firm is to be considered bankrupt according to Swedish legislation.

2.1 PRIOR RESEARCH

Assessing the probability of bankruptcy is not a new research field. There are many researchers that have published articles about it. Some of them are William H. Beaver (1966), Edward I. Altman (1968), James A. Ohlson (1980), Yaw M. Mensah (1984), Andrew W. Lo (1985) and Arshad M. Khan (1985). Research is still conducted within this field. Examples of new researchers are John Stephen Grice and Robert W. Ingram (2001), Stephen A. Hillegeist, Elizabeth K. Keating, Donald P. Cram and Kyle G. Lundstedt (2004), Hyunjoon Kim and Zheng Gu (2006) and Zhongsheng Hua , Yu Wang, Xiaoyan Xu, Bin Zhang and Liang Liang (2007).

In the article "Financial Ratios and the Probabilistic Prediction of Bankruptcy" that was published in 1980, James A. Ohlson has as a purpose to examine which financial ratios are statistically significant for predicting firm bankruptcy. The sample consisted of industrial firms that have been traded at a stock market or over the counter market (OTC) during the fiscal years 1970 -1976. Even though most researchers use multivariate discriminant analysis methods (MDA), James A. Ohlson uses conditional logit analysis methods. He claims that predictors must fulfil some distributional requirements. The two groups (bankrupt and non bankrupt firms) must have the same variance-covariance matrices and the predictors are required to be normally distributed. From this study James A. Ohlson obtains empirical results that shows that size, total liabilities divided by total assets, net income divided by total assets, funds provided by operations divided by total liabilities, working capital divided by total assets and current liabilities divided by current assets are statistically significant for predicting corporate bankruptcy.

A recently conducted study that was published in 2006 is the article "Bankruptcy Prediction: Application of Logit Analysis in Export Credit Risks" written by Li-Chiu Chi and Tseng-Chung Tang. The researchers were interested to study publicly traded firms on Hong Kong's, Japan's, Korea's, Malaysia's, Singapore's, Thailand's and Philippines's capital market during the period 2001-2003. The data consists of 240 firms from which 60 firms have defaulted. Every defaulted firm is matched with three non-defaulted firms based on size, total assets and industry belonging. The empirical results that were obtained from this study indicate that the logit method gives satisfying prediction accuracy and according to the writers the method is robust.

In 1968, Edward I. Altman, in his article “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy” has as purpose to investigate the quality of financial ratios when they are used to assess the potential of bankruptcy. The financial ratios that were used are working capital to total assets, retained earnings to total assets, earnings before interest and taxes to total assets, market value equity to book value of total debt and sales to total assets. The statistical method that is used in order to conduct this research is the multiple discriminant analysis method (MDA). The sample consists of 33 bankrupt publicly traded manufacturers and they are matched with 33 non-bankrupt public traded manufacturers. The multiple discriminant analysis method (MDA) was highly accurate, 94% of the predictions were correct for up to two years ahead of default. Altman noticed that the financial ratios were deteriorating when the bankruptcy date was approaching. It was also observed that there were big changes in the ratios on the third and second year before the year of bankruptcy.

In the article “Logit versus Discriminant Analysis, A Specification Test and Application to Corporate Bankruptcies” which was published 1985, Andrew W. LO compares logit and discriminant analysis methods. The sample that is used in this article consisted of 38 firms that went bankrupt during the fiscal years 1975-1983 which are matched with non-bankrupt firms which belong to the same industry. All firms are obtained from Standard and Poor’s COMPUSTAT Industrial Research File. According to the empirical results, discriminant analysis is better than logit.

One more article that compares logit and multivariate discriminant analysis (MDA) is “Predicting Restaurant Bankruptcy: A Logit Model in Comparison with a Discriminant Model” written by Hyunjoon Kim and Zheng Gu in 2007. The sample consists of 36 U.S. restaurant firms that were publicly traded during the period 1986-1988. The 18 bankrupted restaurants are matched with financially healthy restaurants that had approximately the same amount of assets. The logit method was a little bit better in-the-sample prediction with an accuracy rate of 94% and the multivariate discriminant analysis (MDA) had an accuracy rate of 92%. When the two methods are tested on the out-of-sample data, both resulted with an accuracy rate of 93%. Even though the two methods almost obtain the same accuracy rates, the logit method is preferred because of theoretical reasons. According to Hyunjoon Kim and Zheng Gu, the multivariate discriminant analysis method was conducted without testing for multivariate normality.

On the other hand, Moshe Leshno and Yishay Spector use neural networks in their article “Neural network prediction analysis: The bankruptcy case” that was published in 1996. The purpose of this article is to examine neural network’s ability to predict firm bankruptcy. The sample consists of 44 bankrupted firms that were traded in the New York Stock Exchange and non-bankrupted during 1984-1988. One more requirement was that the firms must have at least 10 million dollars in total assets. Every bankrupt firm is matched with a firm that belonged in the same industry and had approximately the same size of total assets during the last three years before the bankruptcy year. Moshe Leshno and Yishay Spector conclude that neural net models are more accurate than discriminant analysis models.

Wullianallur Raghupathi, Lawrence L. Schkade and Bapi S. Raju in 1991 use neural network methods in their article “A Neural Network Application for Bankruptcy Prediction”. The sample consists of 102 firm and the 51 bankrupted firms were matched with non bankrupted firms which belonged to the same industry and had approximately the same amount of total assets the last three years before the bankruptcy year. All firms were listed at Wall Street Journal Index during the fiscal year 1980-1988 and belonged in a list of deleted firms in the

Moody's Industrial Manual. According to the empirical results of this article, practitioners can not use financial ratios to accurately predict corporate bankruptcy.

One more recently conducted research about corporate bankruptcy which is using neural networks is the article "Comparative Analysis of Artificial Neural Network Models: Application in Bankruptcy Prediction" written of Chris Charalambous, Andreas Charitou and Froso Kaourou in 2001. They have chosen to use an artificial neural network method (ANN) instead of a simple one. The researchers' ambition is to compare the prediction accuracy of artificial neural network methods (ANN) with logistic regression and the feedforward network using BP algorithm. Their data consist of 139 US firms which filed bankruptcy during 1983-1994. The defaulted firms were matched with non-defaulted firms according to the industry belonging, size and bankruptcy year. The empirical results that they obtained indicate that the artificial neural network has higher level of prediction accuracy than the logistic regression method and the feedforward network which is using BP algorithm.

Another study that compares two types of artificial neural networks (ANNs), categorical/instar ANN and probabilistic ANN with the multivariate discriminant analysis (MDA) and logit is the article "A Comparison of the Relative Costs of Financial Distress Models: Artificial Neural Networks, Logit and Multivariate Discriminant Analysis" which was written in 1997 by Harlan L. Etheridge and Ram S. Sriram. The data sample contains 1139 banks from USA during the period 1986-1988 from which 148 banks have defaulted. According to the empirical results the researchers obtain the categorical/instar ANN and the probabilistic ANN have more accurate bankruptcy prediction ability than multivariate discriminant analysis and logit when the prediction is done on a long prediction horizon.

2.2 BANKRUPTCY THEORY

Debt issuers want to be able to quantify the amount of credit risk to which they are exposed. They want to be able to quantify the probability that a borrower will go bankrupt as well. Theories and models have been developed so that practitioners can apply them and obtain the level of creditworthiness of a firm.

2.2.1 CREDIT RISK

Credit risk is the uncertainty that an obligor will not fulfil its obligations (paying back a loan, paying interests, paying suppliers etc.).[4] To assess a firm's credit risk, three factors have to be studied and obtained:[5]

- **Probability to default:** Is the probability that the counterpart will not be able to fulfil its obligations. The probability to default is also called expected default frequency when it is calculated for one year.
- **Recovery rate:** The amount that can be recovered through reconstruction, renegotiation or bankruptcy proceedings.
- **Credit exposure:** Is the amount of any outstanding credit when a firm has filed bankruptcy.

[4] http://www.riskglossary.com/link/credit_risk.htm , 08/04/2007
[5] http://www.riskglossary.com/link/credit_risk.htm , 08/04/2007

2.2.2 Five C's of Credit

Banks' purpose is to provide the borrowers with loans because this is one of the major ways to generate profits.[6] Commercial loan bankers usually use the five C's of credit in order to determine if a borrower should be given or not a loan (Beaulieu, 1996). The five C's of credit are the following (Ibid):

- **Character:** How focused and willing the management of the firm is to repay debt (Beaulieu 1996). Commercial loan bankers examine this by controlling the borrower's credit history.[7] While the commercial loan banker analyses the borrower's character, the knowledge and the experience that the managers of the firm have within the industry is also taken into consideration.[8] Moreover, if the borrowing firm has plans that will enhance the operating and financial strength of the firm within the maturity date then that is an indication that the firm will be able to repay its debt.[9]
- **Capacity:** The management's ability to operate the business with debt in their balance sheet (Beaulieu 1996). A commercial loan banker is analyzing "Capacity" by looking into financial statements and factors like management's experience is taken into consideration (Beaulieu 1996).
- **Capital:** The commercial loan banker can gain all necessary information about available capital that a firm has by looking into the financial statements (Beaulieu 1996). A loan might not get approved by a commercial loan banker if the borrower has not invested personal assets.[10] If a borrower has invested own assets in a firm that indicates that the borrower believes in the firm and that he/she is willing to work in order to succeed.
- **Conditions:** Even if all the criteria that are mentioned above are fulfilled, there is risk that the commercial loan banker will not approve the loan.[11] The lender has to examine the current economic situation and how it will be in the future.[12] Macroeconomic factors like recession and market growth are taken into consideration (Beaulieu 1996). The commercial loan banker will also consider the competition and the development potential of the industry.[13] Not only the industry to which the firm belongs to will be analysed but also industries that might affect the borrower's performance.[14]
- **Collateral:** This is an important determinant, especially if the firm has low quality in "Character, Capacity, Capacity and Conditions" (Beaulieu 1996). If the debt issuer collateralizes some of the borrowers assets then the risk of not getting the loan and the interests paid back is reduced (Beaulieu 1996). On the other hand, the borrower will try to give as less as possible as collateral to the lender because in that way the borrower will not need to have permission from the lender when important decisions has to be made.[15]

---

[6] http://www.sbrn.org/Connections/10_03_Commercial_Lending.htm, 01/05/2007
[7] http://www.bizjournals.com/birmingham/stories/2002/07/01/focus6.html?page=1, 01/05/2007
[8] http://www.ansci.cornell.edu/tmplobs/doc77.pdf, 01/05/2007
[9] http://www.ansci.cornell.edu/tmplobs/doc77.pdf, 01/05/2007
[10] http://www.loanuniverse.com/credit.html, 01/05/2007
[11] http://www.ansci.cornell.edu/tmplobs/doc77.pdf, 01/05/2007
[12] http://www.ansci.cornell.edu/tmplobs/doc77.pdf, 01/05/2007
[13] http://www.loanuniverse.com/credit.html, 01/05/2007
[14] http://www.loanuniverse.com/credit.html, 01/05/2007
[15] http://www.ansci.cornell.edu/tmplobs/doc77.pdf, 01/05/2007

2.2.3 Credit Rating Systems

There are firms that are specialized in assessing the creditworthiness of firms. The most important credit rating agencies are Standard & Poor's and Moody's. As we can see from Exhibit 1, they use similar grading systems in order to give investors an understanding about the creditworthiness of firms (Kim et al 2004).

Standard & Poor's	**Moody's**
AAA	Aaa
AA	Aa
A	A
BBB	Baa
BB	Ba
B	B
CCC	Caaa
CC	Caa
C	Ca
D	C

EXHIBIT 1. Credit ratings where AAA is the best and D/C is the worst

Moody's uses 1, 2 or 3 to indicate if the firm is in the bottom or middle or top of the grade that it is given (Kim et al 2004). On the other hand, Standard & Poor's uses the signs + or – (Ibid).

The credit agency firms might have information about the quality of a firm because firms might give them access to information that is not public (Kisgen 2006). Firms are not always willing to reveal too much insider information to the credit rating agencies because of the risk of revealing valuable information to competitors (Ibid).

There is empirical evidence about Standard & Poor's and Moody's that indicates that if one of those two credit agencies does a change in the rating of a firm or a government then the other credit agency will probably do the same (Guttler et al 2007). It is more likely that they follow each other at a downgrade rather than at an upgrade (Ibid).

2.3 Bankruptcy Regulation

Since the sample of this thesis consists of Swedish corporations, it is necessary to study the bankruptcy regulations that are employed in Sweden in order to understand which firm is considered as bankrupt and under what conditions. The bankruptcy regulations are not going to be examined in detail because it is beyond the purpose of this thesis.

According to the Swedish law, it is not only the borrower that has the right to apply for bankruptcy but also creditors.[16] Borrowers or creditors can file for a bankruptcy by handing in a written application to the district court. [17] But if a creditor requests the bankruptcy of a firm then all necessary information and documents about the claim has to be handed in.[18]

[16] Konkurslag (1987:672), 1 kap 2 §.
[17] Konkurslag (1987:672), 2 kap 1 §.
[18] Konkurslag (1987:672), 2 kap 4 §.

A firm is insolvent when it is not able to fulfil its obligations towards the creditors (pay back loans, interests and suppliers etc.) and the firm is not temporarily into financial distress[19]. But if a creditor has a claim which has been confirmed by the court or the enforcement agency then the creditor will not have any problem to apply for a bankruptcy.[20]

When the bankruptcy application has been accepted then the court has to decide who will be the administrator and how many they will be.[21] The administrator that is recruited has to have all necessary knowledge and experience that is required so that all creditors and employees get as much as possible back.[22] In order to be chosen as administrator you are not allowed to be an employee of the court or to have any relation with the bankrupted firm or with the creditors or with someone else that can affect the administrator's objective work.[23] The owners of the bankrupted firm have the obligation to hand in all the necessary information for which the court, supervisory authority, administrator and supervisors will ask.[24] Afterwards the administrator has to write a report which has to be handed in, no later than six months after the bankruptcy application was accepted, to the court, supervisory authority and to each creditor that request it.[25] The report must contain estate current condition, the reasons why the firm defaulted and a summary of all the assets and claims.[26] The administrator is also responsible for selling off the assets of the bankrupted firm as soon as possible.[27] The assets can be sold in different ways, but selling them off through an auction is preferred if there is no other choice that is better for the estate.[28]

[19] Konkurslag (1987:672), 1 kap 2 §.
[20] Konkurslag (1987:672), 2 kap 6 §.
[21] Konkurslag (1987:672), 7 kap 2 §.
[22] Konkurslag (1987:672), 7 kap 1 §, 1 styck.
[23] Konkurslag (1987:672), 7 kap 1 §, 2 styck.
[24] Konkurslag (1987:672), 6 kap 2 §
[25] Konkurslag (1987:672), 7 kap 15 §
[26] Konkurslag (1987:672), 7 kap 15 §
[27] Konkurslag (1987:672), 8 kap 1 §
[28] Konkurslag (1987:672), 8 kap 6 §

2.4 QUANTITATIVE MODELS

The data that are collected are processed in various statistical and econometric software to create functional and applicable models, for usage in bankruptcy prediction. This study employs the following software for data processing: (i) SAS; (ii) SPSS; (iii) EViews and (iv) NeuroShell. All of these packages are well-renowned and frequently used in the research community. The software is used to create the quantitative model which includes the relevant variables that predicts the probability of default.

2.4.1 MULTIPLE DISCRIMINANT ANALYSIS

The Multiple Discriminant Analysis (MDA) is a multivariate analysis to apportion individuals with several measured characters into one or other of two completely specified population groups (Welch 1939). MDA is a useful categorization tool to apply on the type of problem examined in this study because of its ability to separate objects into two groups, defaulted and non-defaulted, conditional on certain characteristics that they display. The two populations are known to have traits that separate them from each other, for instance the categorization of sexes conditional on the muscle mass measured, or another binary type problem. MDA techniques can also help clarify which variables contribute in making the correct classification, thus MDA have two applications, prediction and description (Afifi et al 1990). In this study, from a credit risk management point of view, the problem could be formulated as defaulted firms falls into populations I and non-defaulted into population II, conditional on the firms financial characteristics. In this case group I firms would be considered as “bad” firms for a credit grantor, as they have a propensity to not fulfil its liability payments, and group II firms would be considered “good”:

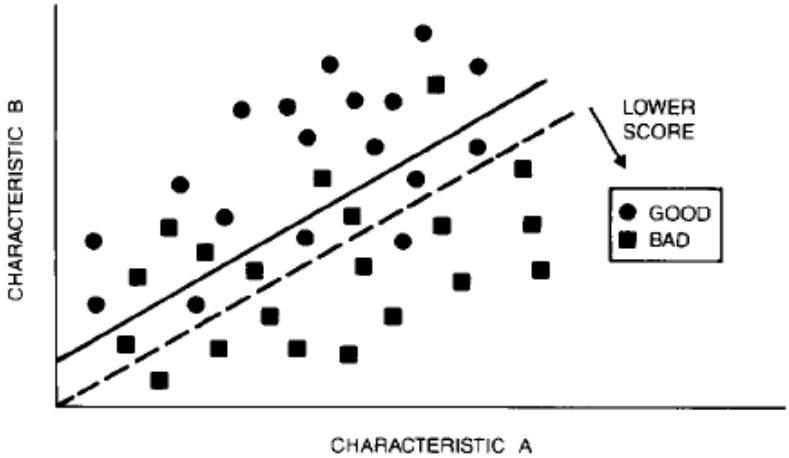

FIGURE 1. Objects fall into different categories conditional on their characteristics (Rosenberg et al 1994).

The different variable values of all sample firms can be scattered in a diagram. Consider a simple bivariate case of only one variable to characterize group I and group II. Plots of the two groups should appear as two separate scatters for each group if the characteristic separates the two groups:

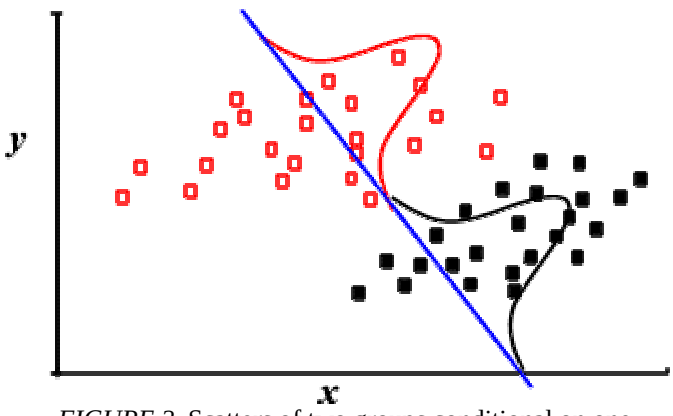

FIGURE 2. Scatters of two groups conditional on one characteristic with the distribution[29].

Of the scatters the probability distributions can be estimated for each population conditional on the characteristic measured. This situation can also be generalized to include several measured variables that are thought to separate the two groups (Afifi et al 1990).

The basic idea is that the distribution of the two populations conditional on certain measured variables differs from each other in terms of their mean values. To obtain the function which most accurately categorizes one object into the correct group the maximum distance between the means of the two probability distributions is calculated (Lachenbruch 1976):

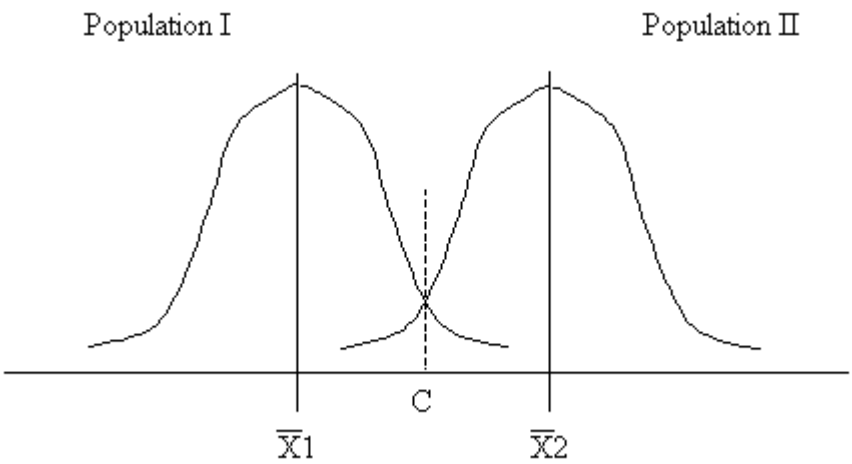

FIGURE 3. The probability distribution of each population and their means, X1 and X2, and cut-off point, C.

It is the distance X1-X2 that is of importance for separation of the two populations since this distance is a measure of how much the populations differ. The distance X1-X2 is called the Mahalanobis distance and can easily be calculated as:

$$D^2 = \frac{(\bar{X}1 - \bar{X}2)^2}{S_X^2} \qquad \text{where } S_X^2 \text{ is the pooled sample variance} \qquad (1)$$

[29] http://149.170.199.144/multivar/dawords.htm

From this simple distance example the analysis can very easily be generalized to include several variables as in the multivariate case, in which case the discriminant function would look like:

$$Z = a_1Y_1 + a_2Y_2 + \ldots + a_iY_i \qquad (2)$$

The multivariate discriminant function contains a set of explanatory variables to explain the population Z, as oppose to the bivariate case. The Mahalanobis distance is discretely transformed to:

$$D^2 = \frac{(\overline{Z}1 - \overline{Z}2)^2}{S_X^2} \qquad (3)$$

There is rarely or never total separation between the two populations that is the tails of the populations respective distributions cross at some point. This leads to a misclassification of some species of the total sample (Lachenbruch 1968). Misclassification of species into other groups than they due to their belonging should have been classified is known as type I and type II errors. Type I errors are in our case group I firms that are misclassified as group II firms and type II errors are group II firms that are misclassified as group I firms. From this logic it follows that there exist an critical point where the two distributions cross and creates a cut-off point between classes. That is population I species should have a discriminant score, or Z-score, that is lower than the cut-off score if accurately classified and population II should have a Z-score that is higher than the cut-off score if accurately classified. The cut-off point is calculated as follows:

$$C = \frac{\overline{Z}1 + \overline{Z}2}{2} \qquad (4)$$

The cut-off point is used to categorize defaulted and non-defaulted firms in to their respective group depending on what Z-score that turns up for each individual firms using the Z-function obtained when processing the data.

2.4.2 Logistic Regression

Logistic regression is a binary response type model and is used whenever an object is to be classified into one of *two* populations (Afifi et al 1990). The regression can either be of a simple kind with one explanatory variable or of a multiple type with several explanatory variables as in this study. Coding of the dependent variable is essential for the logistic regression (logit) and as the name binary implies, the two groups is coded with either a 1 or a 0, making the dependent variable a dummy variable (Cox 1972). The multivariate logit has the following is displayed as follows:

$$Y(1,0) = \alpha + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_iX_i + u \qquad (5)$$

Thus the dependent variable is binary in its nature and can only take on numbers between 0 and 1:

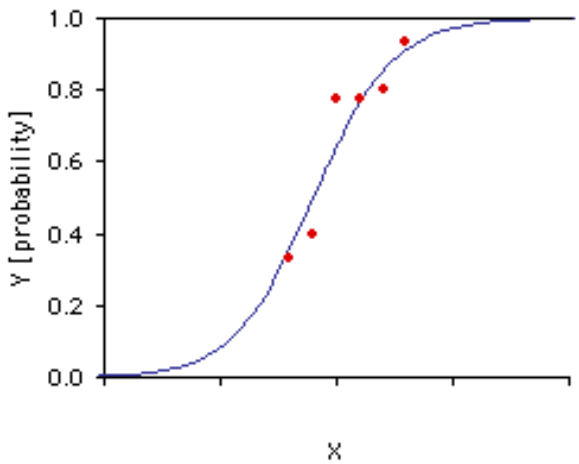

FIGURE 4. A binary response model[30].

As seen in the figure above the dependent variable Y can only take on values between 0 and 1, which makes the logit applicable for binary type problems such as classifying firms into two different groups, defaulted and non-defaulted firms. One main feature of the logit regression and advantage over the MDA is the natural cut-off point that logit gives at 0,5 (Efron 1975). That is firms coded as ones (defaulted firms) should return a number over 0,5 if correctly classified and firms coded as zeros (non-defaulted) should return a number less than 0,5 if accurately classified. Figure 4 displays three firms above the cut-off score and two firms below the cut-off score. If the model in question is a hundred percent accurate in classification the three firms above the 0,5 critical cut-off point should be defaulted firms and the two firms below the critical value should be non-defaulted firms. The number yielded, between 0 and 1, from the logistic regression can thus be seen as a probability of defaulting, because (Cox 1972):

$$\ln\left[\frac{P}{1-P}\right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_i X_i + u \quad (6) \qquad \text{thus} \quad \ln\left[\frac{P}{1-P}\right] = Y$$

and the transformation to probabilities is easily done by:

$$P = \left[\frac{1}{1+e^{-(\alpha+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_i X+u)}}\right] \quad (7) \qquad \text{where} \qquad P \text{ is the probability of defaulting}$$

Because of its binary nature and easy interpretability, logit regression is frequently used in assessing bankruptcy probabilities. There are however special forms of logit regressions which are based on omitting or including explanatory variables depending on what they provide in additional explanatory power looking at the best fitted model according to maximum likelihood. These logistic regression models are known as (Gong 1986; Tegerstedt et al 2005):

(i) Forward Conditional Logistic Regression — Includes variables stepwise conditional on the provision of extra explanatory power in the total model penalized for the additional extra degrees of freedom,

[30] http://faculty.vassar.edu/lowry/lr1.gif

	starting at the variable with the highest explanatory power.
(ii) Backward Conditional Logistic Regression	Omits variables stepwise conditional on the exclusion of minimum explanatory power in the total model penalized for the loss in degrees of freedom, starting with all variables originally specified in the model.

Financial data is seldom normally distributed. The logistic regression however does not require the data to be normally distributed in order to carry out efficient coefficient estimates (Mantel et al 1974). This is a clear advantage with the logistic approach when comparing means, that it does not require exact knowledge of the distribution form (Mantel et al 1974).

2.4.3 Neural Networks

Neural network is a statistical model that is being used within many research areas (Warner et al 1970). According to Warner et al (1970), neural networks were initially used to model the human brain. Afterwards, researchers tried to predict heart problems in patients, to identify underwater sonar contacts, predicting secondary protein structures and predicting stock market performance with the help of neural networks (Warner et al 1970).

Neural network is a nonlinear statistical model which was developed to be used on large samples and with many predictors (Kutner et al 2004). This model is a network of simple units which are called nodes (Warner et al 1970). If it is assumed that there are n inputs ($x = x_1,...,x_n$) then every node can be described by (Leshno et al 1996):

- a threshold value (θ_j)
- a univariate activation function ($s: R \rightarrow R$)
- a vector of weights ($w_j = w_1,\ldots,w_n$)

The nodes consist of input nodes that are the base layer and the output nodes that are the top base (Bortiz et al 1995). Moreover there is at least one hidden layer that is between the input and output layer (Ibid). If the weights of the input nodes are positive then they are excitatory otherwise they are inhibitory (Ibid).

Neural networks' structure can be either feedforward or feedback (Ibid). The feedforward method is also called single-hidden-layer (Kutner et al 2004). If the feedforward method is chosen to be used in a study then all inputs are received from previous layers (Bortiz et al 1995). That means that all information moves from the input nodes, to the hidden layers and finally to the output layers (Ibid). The functional expression for a feedforward neural network is (Leshno et al 1996):

$$f(x) = \sum_{j=1}^{k} \beta_j s(w_j x - \theta_j) = \sum_{j=1}^{k} \beta_j s\left(\sum_{i=1}^{n} w_{ji} x_i - \theta_j \right) \qquad (8)$$

where

x = Input vector
β = Weight vector associated with the single output unit
k = Denotes the number of processing units in the hidden layer

As can be seen in Figure 1 all information has a forward direction (Boritz et al 1995).

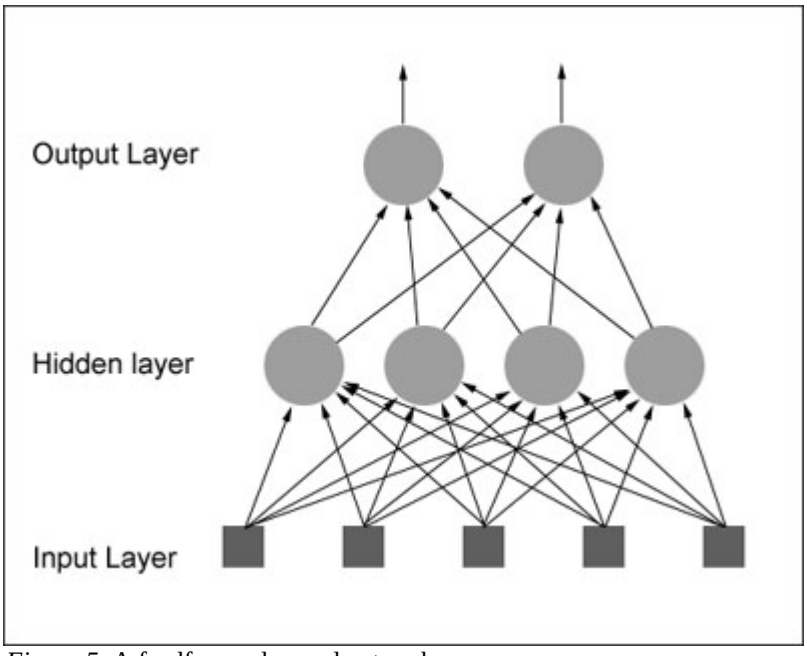

Figure 5. A feedforward neural network.

On the other hand, if a feedback neural network is used then inputs can be received from all layers which means that the information flow does not have a certain structure (Ibid) (see Figure 2).

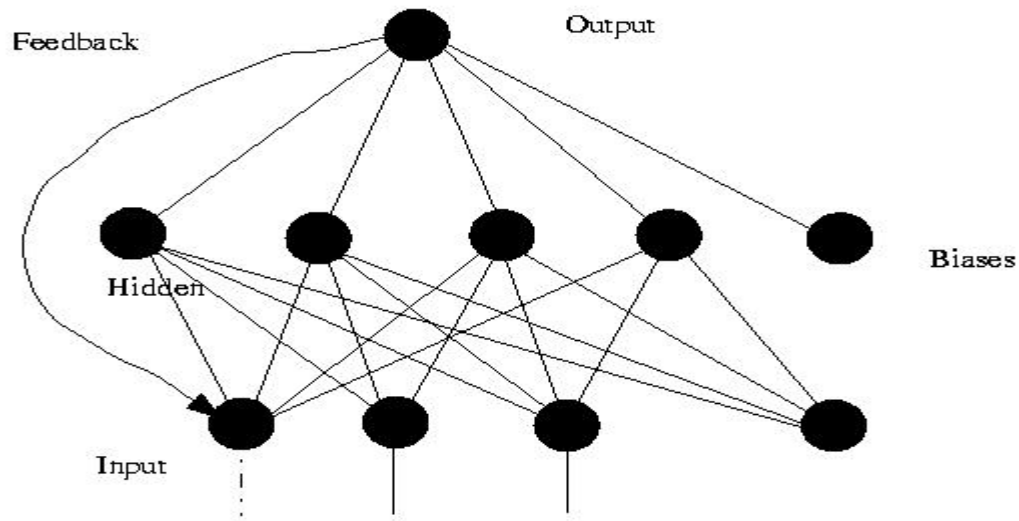

Figure 6. Feedback neural network (Boritz et al 1995)

The feedback neural network has the following functional form (Salam et al 1993):

$$C_i \frac{du_i}{dt} = \sum_j T_{ij}^2 S_j(u_j) + I_i - \rho_i^{-1} u_i \qquad i = 1, \ldots, n. \tag{9}$$

where

S_j = Is a bounded differentiable monotone increasing function. Neuron i is connected to neuron j via the weight T_{ij}^2

I_i = Is the external input supplied to node I $\qquad$ C_i and ρ_i are constants

The probabilistic neural network is based on the Bayesian decision making theory which estimates the probability density in data space. This type of neural network has four layers. The first layer contains the input data and the second layer contains patterns of the units. At the second layer it is also estimated the contribution of each pattern to the density function. The third layer is also called the summation layer and the density estimation for each pattern is summarized. Finally, at the last layer the probabilistic neural network makes a decision based on the Bayesian decision making theory. (Yang 1999)

The function that the probabilistic neural network is using is (Ibid):

$$p(x|c) = \sum_{m=1}^{M_c} \overline{\omega}_m^c \, p(x|u_m^c) = \sum_{m=1}^{M_c} \overline{\omega}_m^c \left(\frac{\beta_m^c}{2\pi} \right)^{d/2} \cdot \exp\left(-\frac{\beta_m^c \left| x - u_m^c \right|^2}{2} \right) \tag{10}$$

where

M_c = The number of pattern units

$\overline{\omega}_m^c$ = Is the mixing coefficient of the mth pattern units of the cth class

u_m^c = Is the center of the mth pattern units of the cth class

β_m^c = Is the smoothing parameter corresponding to the mth pattern units of the cth class

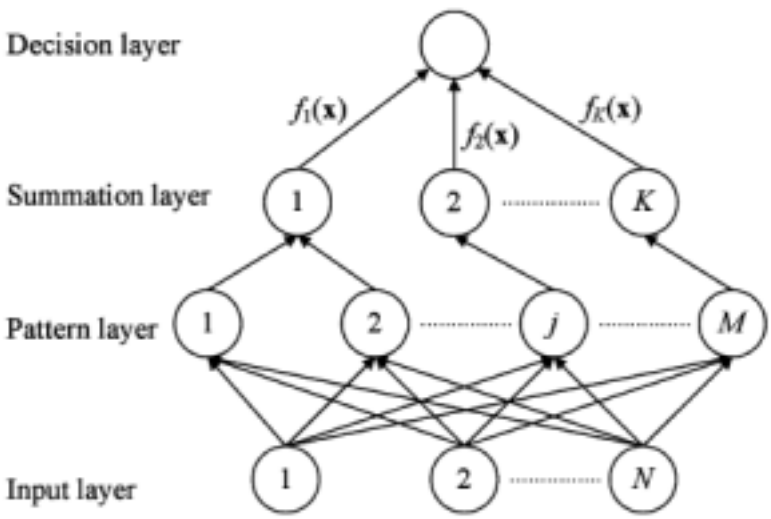

Figure 7. The architecture for a probabilistic neural network (Gaganis et al 2007)

The conduction of a neural network can be divided into two parts, the “learning” and the “testing” part (Boritz et al 1995). The purpose of the learning part is that the neural network will learn how to solve a problem or to find structures based on information that is given by the input data (Ibid). There are two ways to proceed with the learning part, the unsupervised and the supervised learning (Ibid). In the unsupervised learning method, the network is provided with data and no response is known by it (Ibid). The network processes the data and an algorithm is used to classify the data (Ibid). On the other hand, the supervised learning method is provided with data with correct response (Ibid). The network uses the information provided from the data and creates connections weights for every node (Ibid).

Researchers usually conduct the training part with the help of a method that is called Back-Propagation which is a feedforward network (Ibid). This method adjusts the input weights in order to minimise the difference between the output and the target (Ibid). There are different versions of Back-Propagation, the functional link Back-propagation with slines, pruned Back-Propagation and predictive cumulative Back-Propagation (Ibid). The functional link Back-Propagation with sines use an algorithm to adjust the input weights and adds more nodes in the input layer (Ibid). By adding additional nodes, the learning rate of this method is increasing (Ibid). The pruned Back-Propagation method does not take into consideration any connections that have little influence on the output (Ibid). If the complexity of neural networks increase then the generalization ability declines (Ibid). Finally, in the predictive cumulative Back-Propagation a linear transfer connection is used for the input and output layer and a sigmoid transfer function (Ibid). The advantage with this method is that the learning time can decrease with 50% (Ibid). In order to decrease the duration, this method uses a rule that is called the cumulative generalized delta rule (Ibid). The principle behind this rule is that after two data units have been presented to the network then the weights are updated (Ibid).

Neural networks can be found in different versions. Artificial neural network (categorical learning/instar neural network and probabilistic neural network are two different versions of artificial neural networks) and auto-associate neural network are some examples of neural networks.

One advantage that neural network has is that it does not require the assumptions of a linear regression (Kutner et al 2004). The assumptions that the residuals have to be independent, normally distributed and have constant variance are not necessary when neural networks is used (Ibid). Moreover, neural networks is flexible because it can be used in order to model sufficient data, curvatures, interactions, plateus and step functions (Ibid).

One problem that neural networks face is that there are more healthy firms than bankrupted and as a result the data imbalance problem is created (Baek et al 2003). For this reason Baek et al (2003) in their study they compared an auto-associate neural network (AANN) with a two class neural network. The empirical results that were obtained indicate that the auto-associate neural network (AANN) has better bankruptcy accuracy prediction. Moreover, some disadvantages with neural networks are that it requires a large data sample and the parameters can be difficult to interpret (Kutner et al 2004). Finally, identification of outliers and significance testing of the influencing power of the predictors are not available in the most versions of neural networks (Ibid).

CHAPTER 3
METHODOLOGY

This part of the thesis involves giving a detailed description of the research method employed. This chapter gives an extensive overview of the data gathering process and the characteristics of the firm that are included in the study. Further, the different mathematical models are presented.

3.1 SOURCES OF INFORMATION

In this study the focus was to obtain the financial data from *one* source only. This is to ensure reliability of the data that is included in this study. For this purpose the Swedish information database, Affärsdata, was selected. All *financial* information is extracted from this source of information. Having multiple databases for collecting financial data might cause problems in the accuracy of the numbers, since different databases might employ different methods to calculate different items in the balance sheet, income statement and pre-specified ratios.

Following databases are used in this study:

DATABASE	*DESCRIPTION*
Affärsdata	Extraction and gathering of included firms matched firms and their respective financial information.
ELIN	Searching and collection of relevant articles and search for prior research and theory of importance in this study.
SCB	Gathering of relevant bankruptcy statistics for Sweden.
LIBRIS	Used for search of relevant literature.

The data collected from Affärsdata are the main data of importance in this study. The data collected from that source come in a raw-data form, that is balance sheet items and income stat items are presented. The ratios are then calculated by us; For example two items in the balance sheet like long term debt and equity are calculated by us as a form of leverage ratio (long term debt to equity). Some of the ratios are however given directly by the database, such as change in sales and change in total assets. For a complete overview of the ratios and calculated items see Table 4.

3.2 CRITICISM OF SOURCES

All of the sources used must be considered as reliable as they are national databases for information. However second hand information are used as financial information, which means that no direct financial information concerning the sample firm is collected from their own financial statements. Only financial information from Affärsdata is extracted, treated and processed. This fact might cause a problem with biased data.

Preferably, in using information such as books and articles, first hand well-renowned articles and publishing houses are consulted. Even though this principle is employed there is always a question about articles and books objectiveness of their own research and caution must be applied. Also, newer articles used in this thesis might lack a well-renowned publisher which invites scrutiny.

Statistical databases such as SCB are considered highly reliable sources and data extracted from here are considered accurate. Only human error in processing is the only realistic flaws that might influence the data.

3.3 DATA COLLECTION

To collect the necessary data for this study a set of defaulted firms are gathered that meet our requirements (see section 1.4 LIMITATIONS). The Swedish information database, Affärsdata, is used for this purpose and returned 115 firms matching our specifications:

Specifications
(i) Number of employees; At least 30
(ii) Bankruptcy year; 2004-2007
(iii) Type of firm; Swedish Corporation (Aktiebolag, AB)
(iv) Municipal court code for bankruptcy[31]: 20, 21, 22, 25

Out of the 115 hits two financial firms is excluded and a final sample of 113 (see Appendix III) is obtained. After obtaining the defaulted firms the counter group of non-defaulted firms are collected. These firms, which constitute the contra or counter sample are matched by a pair-wise industry principle. That is each of the defaulted firms is allotted a firm belonging to the same industry according to a five digit Swedish industry classification (SNI-code). The motivation behind a pair-wise industry matching is the fact that different industries might over time have different possibilities of obtaining capital. Prior research have pointed to the following relationship between the probability of bankruptcy and firm leverage, i.e. $PD=f(D/E)$[32] (Altman 1968), that is the more levered a firm is the closer they get to the point of bankruptcy, as financial distress cost kicks in with increased levels of debt. Different industries have different means of obtaining capital at different points in time. One industry can differ severely from another by its ability to obtain liability free capital. As an example can be mentioned the willingness of investors to grant equity capital to firms in IT related industries during the late 90's. This fact can potentially cause biasness due to different industries being able to manipulate leverage levels easier and thus potentially lower its probability of defaulting. This matching principle is common practice in the research

[31] Court codes: 20-Bankruptcy started; 21-Bankruptcy closed; 22-Bankruptcy closed with additional surplus; 25-Bankruptcy continues.
[32] PD is the probability of defaulting

community and also employed here, however problems with industry matching have been reported. Studies with model-accuracies of up to 87% without using industry type matching exist (Sheppard 1994).

TABLE 1

INDUSTRY CLASSIFICATION OF TOTAL SAMPLE

Industry	*Defaulted*	*Non-defaulted*	*Total*	*SNI-Code*
Forestry	1	1	2	**02**
Foods	2	2	4	**15**
Textiles	2	2	4	**17**
Paper	3	3	6	**21**
Graphics & Publishing	3	3	6	**22**
Oil	1	1	2	**23**
Polymeric	1	1	2	**25**
Metals	2	2	4	**27**
Metal Products	2	2	4	**28**
Machinery	5	5	10	**29**
Electronics	1	1	2	**31**
Instrumental & Optics	1	1	2	**33**
Transportation	1	1	2	**35**
Construction	14	14	28	**45**
Automotive Trade	3	3	6	**50**
Wholesale	4	4	8	**51**
Retail	5	5	10	**52**
Hotel & Restaurant	2	2	4	**55**
Land Transportation	2	2	4	**60**
Airlines	1	1	2	**62**
Transport Services	8	8	16	**63**
Telecom	3	3	6	**64**
Real Estate	6	6	12	**70**
Rental Services	1	1	2	**71**
IT & Computer	4	4	8	**72**
Research & Development	2	2	4	**73**
Outsourcing Services	21	21	42	**74**
Educational	1	1	2	**80**
Healthcare	3	3	6	**85**
Other Manufacturing	4	4	8	**36**
Other Services	4	4	8	**93**
TOTAL	*113*	*113*	*226*	

TABLE 1. Industry classifications of the sample firms according to their respective aggregated two-digit SNI-code. A total of 31 different industries are represented in this study.

The final sample consists of 226 firms, 113 from each group. From this data 23 firms from each of the two groups were additionally excluded from the total sample making them a holdout sample. The holdout sample is a random sample that consists of firms from various industries, bankruptcy years and number of employees.

As seen in Table 1 a variety of industries are represented. There is a slight over-representation of firms in the outsourcing industry (18,6% of total sample) and construction companies (12,4% of total sample). A total number of 31 industries are represented in the study.

The total sample consists of industries from a multitude of industry classes, ages and size (see figure 8 and 9; see also Appendix I and II for further information). No discrimination is made according to size or age of the studied firms in the matching process, which is because of the potential of these variables to explain the probability of defaulting.

FIGURE 8

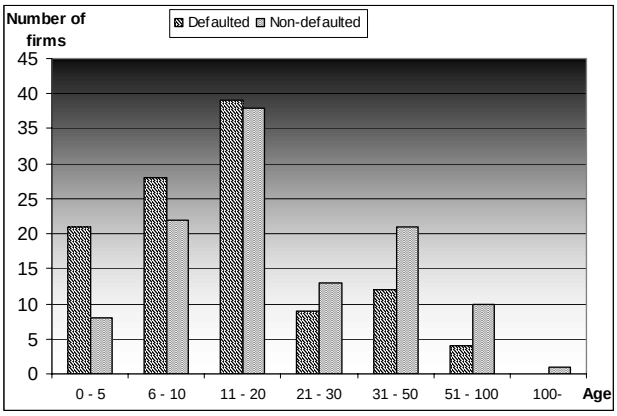

Figure 8. Number of firms in a specific age interval.

FIGURE 9

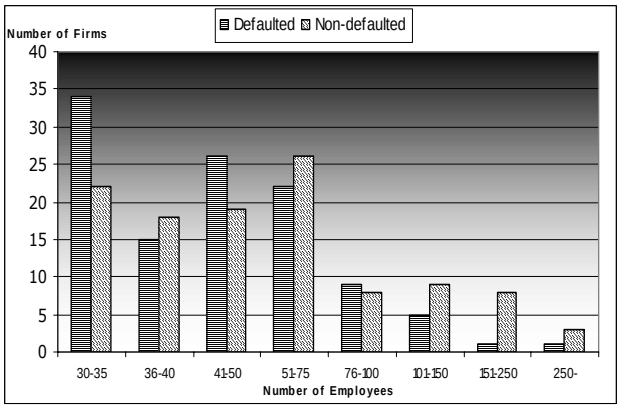

Figure 9. Firm categorized by number of employees

The sample firms are defaulted and non-defaulted firms from the years 2004-2007. According to figure 10 a concentration of bankruptcies of firms included in the study is found in the years of 2005 and 2006.

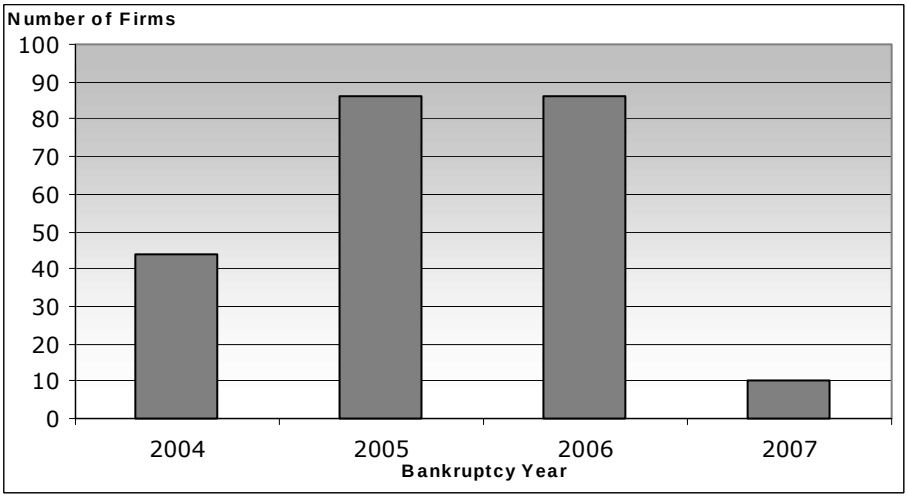

Figure 10. Number of firms in the respective bankruptcy year

The process collecting the data involves taking the necessary figures from the respective firm's financial statements (balance sheet and income statement) one year prior to bankruptcy. That is the figures from the last publicly available statements are considered in this study. Group II firms (pair-wise matched sample) financial information is extracted according to the same principle as for group I firms (defaulted firms) conditional on the point of bankruptcy for the group I firm. This means the following; Firm ZZ defaulted in 2005 and is classified as wholesale firm (two digit SNI-code is 51XXX) and the last publicly available financial information is extracted. Firm YY is the appointed pair-wise match according to the 5-digit SNI-code 51432 (the same as for firm ZZ). The financial statement of firm YY is extracted according to what year firm ZZ defaulted and which point in time the financial information is extracted. Both of the firms would have their financial information extracted from 2004 in this case.

An important criterion for the group II firms is that they have no documented history of bankruptcies or filing for bankruptcy according to the Swedish municipal court orders 20, 21, 22, 25 through history up to present date.

A number of financial ratios and other characteristics of firms are calculated on the sample firms. This is done because of the belief that bankruptcies can be distinguished by looking at the firms financial health. The financial soundness of a firm is believed to differ from a firm that is defaulted rather than an ongoing concern firm in the sense that the economics of the firm is taken a turn for the worse. This notion is what drives bankruptcy studies to examine firm characteristics as the primary predictor of whether or not the firm will live on.

The quantitative study is conducted by calculating 32 different financial ratios and other firm characteristics that has been used in prior studies. The ratios are calculated on all 226 firms, which produce a total of 5760 data-points in-sample and 1472 data-points out-of-sample. This data is processed by three statistical and econometric models, namely: (i) Multiple Discriminant Analysis; (ii) Logistic Regression and; (iii) Neural Networks. All of these models have frequently been used as application tools for bankruptcy studies.

The following is a table over the selected ratios that are included in this study along with the empirical support:

TABLE 2

Measured ratios in the study

	VARIABLE	EMPIRICAL SUPPORT
I. CAPITAL STRUCTURE		
1. Long Term Debt to Equity	(X20)	Arshad 1985
2. Long Term Debt to Total Asset	(X21)	Beaver 1966; Ohlson 1980
3. Short and Long Term Debt to Equity	(X23)	Laitinen et al 2000; Chi et al 2006
4. Short & Long Term Debt to Tot. Assets	(X22)	Beaver 1966; Laitinen et al 2000
5. Cash-Holdings to Total Assets	(X17)	Dugan et al 1989; Laitinen et al 2000
6. Equity Share of Total Assets (Solvency)	(X5)	Cielen et al 2004; Pompe et al 2005
7. Current Liabilities to Total Assets	(X32)	Lo 1985; Spanos et al 2001
II. LIQUIDITY RATIOS		
1. Quick Ratio	(X11)	Cielen et al 2004; Arshad 1985
2. Cash to Long & Short Term Debt	(X27)	Dugan et al 1989; Hua et al 2007
III. PROFITABILITY RATIOS		
1. Return on Equity	(X6)	Pompe et al 2005; Arshad 1985
2. Return on Assets	(X7)	Becchetti et al 2003; Arshad 1985
3. Net Profit Margin	(X10)	Beaver 1966; Kim et al 2006
4. Gross Profit Margin	(X9)	Cielen et al 2004; Kim et al 2006
5. Profit Margin	(X8)	Becchetti et al 2003; Pompe et al 2005
IV. CASHFLOW AND EFFICIENCY RATIOS		
1. Cashflow to Sales	(X25)	Beaver 1966
2. Cashflow to Total Assets	(X18)	Young et al 2005; Pompe et al 2005
3. Cashflow to long & short term debt	(X27)	Beaver 1966
4. Sales to Total Assets	(X1)	Altman 1968; Raghupathi et al 1991
5. EBITA to Total Assets	(X2)	Altman 1968; Young et al 2005;
6. Retained Earnings to Total Assets	(X3)	Altman 1968
7. Sales per Employee	(X12)	
8. Salary Expenses per Employee	(X13)	
9. EBITA per Employee	(X14)	
10. Financial Expenses to Sales	(X24)	Becchetti et al 2003
V. DEVELOPMENT AND FIRM SPECIFICS		
1. Change in Sales from Last Year	(X15)	Chi et al 2006
2. Change in Total Assets from Last Year	(X16)	Hua et al 2007
3. Firms Age	(X4)	Berg 2005; Chi et al 2006
4. Firm Size (Natural Logarithm of Sales)	(X19)	Sudheer et al 2004; Chi et al 2006
VI. TURNOVER RATIOS		
1. Accounts Receivable to Sales	(X28)	Beaver 1966
2. Accounts Receivable to Total Assets	(X29)	
3. Inventory to Sales	(X30)	Beaver 1966; Young et al 2005
4. Inventory to Total Assets	(X31)	Cielen et al 2004

TABLE 2. An overview over the ratios that is assessed in this study and the empirical support in prior research.

3.4 DATA PROCESSING

The data processing involves computer software to assess the input to create relevant models that can then be subjected to test of robustness. Input data are used to train the model and give the model its character. In the case of the discriminant analysis and logistic regression the data used to most accurately categorize bankruptcy is given by:

(a) A discriminant function; $Z_{swe} = a_1Y_1 + a_2Y_2 + \ldots + a_iY_i$ where Z is a number

(b) A logistic function; $P = \left[\frac{1}{1+e^{-(\alpha+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_i X_i+u)}} \right]$ where P is the probability

Subsequently the two functions can easily be applied to predict bankruptcy on any dataset not used to train the model by plugging in the variables yielded by each model to the out-of-sample-set. In the case of the neural network the complex nature of the system sets restrictions on how easy the model is to apply without the help of computer software. Therefore the training of a network is only manageable by computer software and thus stored in the software. This trained neural network can then easily be applied on an out-of-sample set of data by using the software.

In order to find the best predictive model a number of model types are used for the three different approaches:

(I) Discriminant Analysis
- (i) Standard type employed

(II) Logistic Regression
- (i) Standard type
  - Including only variables significant on the 90% level.
  - Including only variables significant on the 80% level.
- (ii) Forward Conditional
- (iii) Backward Conditional

(III) Neural Networks
- (i) Feedforward
- (ii) Backward
- (iii) Probabilistic

In assessing each of the models robustness and predictability both the in-sample accuracy and the out-of-sample accuracy is evaluated. Before development of the actual models variable testing to detect heteroskedasticity, multicollinearity and normality is conducted.

3.4.1 HETEROSKEDASTICITY

Both the discriminant model and the logit model are tested for heteroskedasticity. Heteroskedasticity is present if the error term does not have constant variance (Anscombe 1967). Heteroskedasticity is detected using the White probability test indicator, based on a heteroskedasticity-consistent covariance estimator by Halbert White (1980). If heteroskedasticity is found among the variables this is adjusted for in the data processing either by hand or by available applications in the computer software. With the presence of heteroskedasticity faulty inferences can be drawn when testing the statistical hypotheses (White 1980). No adjustments regarding heteroskedasticity for neural networks are required.

3.4.2 MULTICOLLINEARITY

The problem of multicollinearity relates to the explanatory variables being highly correlated with each other which may lead to imprecision in estimation of the regression coefficients (Harvey 1977). This problem is prevalent in modelling financial data since the variables closely relate to the same raw-data set, the balance sheet and the income statement (Tucker 1996). Multicollinearity is simply remedied by adding additional data or omitting one of the two correlated variables.

In the logistic regression the problem with multicollinearity is most likely to exist and measures to deal with the problem is taken in this study by omitting on of the two or more correlated explanatory variables from the model (see Appendix V for a complete variable correlation matrix). The model containing only uncorrelated variables will then be processed. In the case of the multiple discriminant analysis Eisenbeis (1977) claims that "multi-collinearity is not a problem in MDA if classification accuracy is the objective". Since the purpose of our study is just that, accurately classifying defaulted and healthy firms, no action is taken regarding multicollinearity in the MDA model. Further, according to Tucker (1996) multicollinearity is not likely to be a problem in neural networking and thus disregarded in this study.

3.4.3 NORMALITY ASSUMPTION

The assumption of normality states that the examined variables follow a certain pre-known distribution, the normal distribution. However, financial data rarely displays this type of characteristics but rather skewed (Barnes 1982). Therefore the normality assumption is unlikely to hold when dealing with financial type data.

A test is carried out on the distribution of the variables to check for normality, and the Jarque-Bera test is applied for this matter. Should non-normality exist in our data-set this is adjusted for by relaxing the normality assumption in the MDA model, which might cause loss of model power (see further section 3.5 RELIABILITY AND VALIDITY). According to Mantel et al (1974) the distribution form for the logistic function is not required to know, thus any adjustments according to above stated will not be done in this case. Furthermore, a strength of neural networks is the lack of variable distribution requirements (Tucker 1996), thus no adjustments according to above stated is neither done on the neural networks.

3.5 MODELLING APPROACH

This section describes the systematic modelling of prediction models for bankruptcy applications of MDA, Logit and ANN implemented in this study. The modelling is largely based on the methodological approach suggested by Tucker (1996) and Joy et al (1975).

3.5.1 MULTIPLE DISCRIMINANT ANALYSIS

Model creation and testing
The discriminant procedure involves encoding the two groups for the specific classification problem, defaulted firms and non-defaulted firms. Firms that are defaulted are encoded as ones (1) and non-defaulted firms are encoded as twos (2). The set of variables examined are then processed by an MDA procedure using the Mahalanobis distance as the omission criteria of variables that do not contribute to the model efficiency for the model creation. Of this follows a set of variables with their respective coefficients to give the discriminant or Z-score of each firm. The cut-off score is calculated and used as the classification criteria.

This Z-score is applied on the holdout sample by calculating the respective Z-scores of the out-of-sample firms. Once again the cut-off score is used as a critical point for classification. Note that for all the three modelling procedures (MDA, logit and ANN) the model creation sample and holdout sample consist of the same firms respectively. The model creation and the model testing procedure, which involves calculating the Z-score for the sample firms, is done in EXCEL.

3.5.2 LOGISTIC REGRESSION

(i) Variable specification
The logistic regression is of dichotomous nature, with the dependent left-hand side variable being binary (1 or 0). The left-hand side predictor variables are selected by the relevancy for bankruptcy prediction studying previous research. The predictor variables are generally of absolute nature (e.g. firm age, firm size etc) or ratios (e.g. return on assets, solvency etc). The selection of the predictor variables has implications for their respective distribution form and might invalidate the conventional regression assumptions. However, logistic regression is less sensitive to these assumptions and knowledge of the distribution is not critical. Thus, no action concerning this is taken to deal with the issue.

(ii) The problem of missing values
Financial data sprung from information databases tend to quite often have missing values. Tucker (1996) suggests a number of solutions: (a) Delete variables which contain missing values; (b) Fill in the gap with the mean of the variable in question across the sample; (c) Fill in the gap with a random value from the variable's distribution. In this study suggestion a) is implemented, because of the slight manipulation of data for the other two approaches. This resulted in omission of one variable of originally 33 examined.

(iii) Model development process
In this stage the selection of modelling method comes in and can generally be said to be of two natures; forward development (adding variables in a stepwise manor) and backward development (omitting variables in a stepwise manor). This study implements three type of

logit models: (a) Manually adjusted logit only including variables significant at certain threshold levels; (b) Forward conditional and; (c) Backward conditional. Tucker (1996) claim however that a backward approach would be preferable since this approach suffer less from omitted variable bias in the estimators. The model development in approach a) is based on the Wald statistic significance at the 80% and 90% as threshold values for inclusion of variables.

(iv) Model performance and holdout testing
Each of the obtained models will be tested according to their respective classification accuracy. That is testing the in-sample model how accurate it is in categorizing firms as either defaulted or non-defaulted depending on their actual group belonging.

Further, the developed models are tested with a out-of-sample group of firms containing 20% of total original sample of 226 firms. Thus a total of 180 firms is used for model creation and 46 firms are used for model validation.

3.5.3 NEURAL NETWORKS

(i) Variable specification
The task of specifying the variables for the neural network is much done in the same manor as for the two previous models. It is argued in research that two different pre-processing of the data can be employed: Data truncation and simply leaving the data as it is. The data truncation involves treatment of the data set by omitting outliers and extreme values to smooth the data. However we argue that this approach is faulty for the following reasons: (a) Neural networks are designed to assign weights to variables depending on their respective relevant contribution, thus the network will assign less weight to such variables that are of outlier nature; (b) Neural network are designed to find patterns in the data, and omission of certain data points might falsely create certain data patterns; (c) Omission of variables might be considered as manipulating the data in favour of research efficiency, and furthermore omitting "un-wanted" variables might leave out actual real world conditions and, (d) Neural networks can easily train upon distributions that are of non-normal shape and as result Tucker (1996) argue that the data should be kept in its raw form. In this study the data is left un-manipulated according to above stated argumentation.

(ii) Model training
The training of the network is crucial for accurate model specification. The process of training the network involves assigning the correct structure for the network conditional on the given data input. Thus this process is the actual model building. There is no practical use in trying to find a function to describe the data, since this is beyond the comprehension of the human brain. Rather it is preferable to let the trained network of complex structure be stored in the software for further application on other data sets such as a holdout sample. In this study the trained network is stored in the computer software and no further presentation of a comprehensible functional mathematical form is available, rather than just presenting the results that the NeuroShell software provides.

In the training process the task of assigning number of hidden layers and session time (training time) is left to the researcher. The number of hidden layers assigned to train the network is defaulted at 30 hidden layers. However threshold values of the number of hidden layers are tested at 5, 10, 15, 20, 25, 35 and 40 as well. If the network performance does not improve with the variation of the number of hidden layers the default number is selected and thus the others are disregarded. The training time is also a crucial task in the neural network modelling since overtraining the network might make the network less efficient. However

there is no clear guide of when to stop training the network. One suggestion is based on when the root mean square error (RMSE) converges to a minimum and reaches a “stick” point. Thus the network training time in this study is conducted by visually checking when the training process reaches its “stick” point based on the RMSE, which at this point the error term is minimized and the network optimized. According to the NeuroShell users guide usually no more than 20 000 epochs (an epoch is an entire pass through all of the training patterns) are needed to minimize RMSE. When the network RMSE converges to a minimum it is manually stopped and considered fully trained.

(iii) Model Testing

The network testing involves applying the holdout sample firms to the trained network and thus evaluating the predictability of the network. Should the network consistently perform bad in the testing phase then retraining the network using other parameter settings might be useful as remedy to the problem. Validating the model is done in a similar fashion as for the two other approaches, however model testing can not be done manually by calculating a “score”. Thus model validation is simply done by testing the network on the given holdout sample using the computer software. The trained network is stored and can also be applied on other out of sample datasets.

3.6 RELIABILITY AND VALIDITY

To ensure quality in our work as well as accuracy and to ensure that a replication of this study would yield the same results some precautionary measures are taken:

To ensure *reliability*

- Testing the robustness of the model using a holdout sample which is excluded from the total original sample, thus not a part of specifying the model itself. This is to ensure model reliability.
- Applying the same process of data gathering for all firms and using only one database in collecting the quantitative data.
- Using a standardized spreadsheet with formula function for calculations of all firm's ratios and do random manual recalculations of the ratios.
- Apply well-established statistical- and econometric models to process the data in order to ensure that no quantitative flaws in the model-building are present. Applying the models correctly is of great importance and a standardized processing of data is applied, which involves: (a) Checking the data for flaws; (b) Omit variables that interfere with the model building and omit variables that are multicolinearized in the case of the logistic regression; (c) Running the data in a systematic manor. This approach is important to ensure the test-retest-reliability and the objectiveness.

To ensure *validity*

- Since bankruptcy probabilities are what is measured the incorporation of as many variables that potentially affect the bankruptcy as possible is important. This study incorporates 32 variables which empirically are proven to have an impact on bankruptcy.
- The choice of method is commonly applied in this area of research and well documented in prior research, this involves both data collection and data processing.
- In some research of this kind the assumption of normality is taken for granted. However in this study we check variables with both a parametric test (such as t-test) that assumes normality and a non-parametric test (such as Mann-Whitney U-test) which doesn't assume normality. This is conducted to examine significant difference between firms in group I and firms in group II conditional on each financial ratio. A loss in power of the model might be expected due to the use of more unrestricted non-parametric test.

CHAPTER 4
RESULTS

In the result chapter, extensive information about the obtained results is presented. In a systematic fashion the results are reproduced by initial variable analysis and further each of the model used are presented. The model results are given by in-sample accuracies and holdout testing for each model, so that the reader easy can access the individual model performance. Finally, this section provides a summary table for all the key results of each model giving the reader the chance to get a clear overview of the results.

4.1 DESCRIPTIVE STATISTICS

The descriptive statistics and the variable testing is conducted on the regards of the model development, thus only the in-sample firm characteristics are evaluated. Holdout firms do not have anything to do with model specification thus not of importance for this matter.

Information of the sample firms is compiled and the variables are evaluated in a systematic fashion. No adjustments of extreme values are done in this study based on the notion that this approach is “fact-of-life” manipulation (see further section “3.5.3 Neural Networks”). Table 3 shows a compilation of the descriptive statistics over all included variables for the in-sample firms. The results point to the existence of possible outliers in variables X_6, X_{12}, X_{13}, X_{14}, X_{20} and $X_{23.}$ Although no negligence of the variables is done due to them being a part of the actual data set, extreme values could be a problem for variable interpretation. If the final models should include some of the variables which display possible traces of extreme value, caution must be implemented in drawing inferences of the model’s accuracy capability.

TABLE 3

DESCRIPTIVE STATISTICS OVER THE VARIABLES

Variable	*Min*	*Max*	*Mean*	*Median*	*Std. dev.*
X1	0,18459	17,570	2,9599	2,3926	2,1036
X2	-2,2189	3,8346	0,0274	0,0747	0,4117
X3	-5,7449	0,7400	0,0132	0,0762	0,5217
X4	1,0000	101,00	20,310	14,000	18,293
X5	-1,9517	0,8840	0,1557	0,1782	0,3544
X6	-1100,9	15,565	-6,215	0,1145	73,672
X7	-3,0442	0,6311	-0,051	0,0392	0,3836
X8	-2,6436	0,4041	-0,030	0,0164	0,2903
X9	-2,6436	0,4041	-0,034	0,0147	0,2898
X10	-2,6490	0,3159	-0,045	0,0077	0,2942
X11	0,0000	10,578	1,0670	0,9328	0,9983
X12	0,0000	20392	1530,5	1022,5	1965,9
X13	0,0000	759,96	274,58	270,44	107,52
X14	-2163,8	626,26	-28,00	12,535	298,50
X15	-0,6506	18,804	0,3572	0,0609	1,5820
X16	-0,8990	158,17	0,8616	0,0444	10,528
X17	0,0000	0,7178	0,0979	0,0280	0,1508
X18	-1,7635	3,6903	0,0689	0,0719	0,4045
X19	7,0139	14,994	10,878	10,761	1,0935
X20	-39,782	1251,6	8,7511	0,1311	83,995
X21	0,0000	1,1833	0,1676	0,1088	0,1921
X22	0,0707	2,9517	0,8296	0,8121	0,3652
X23	-115,83	1876,7	21,829	3,7354	130,81
X24	0,0000	0,6179	0,0145	0,0063	0,0450
X25	-2,3728	1,8339	0,0148	0,0323	0,2754
X26	-3,3363	7,3943	0,1344	0,0932	0,6545
X27	0,0000	8,9202	0,2198	0,0289	0,7357
X28	0,0000	0,8022	0,1147	0,1091	0,0965
X29	0,0000	0,8711	0,2753	0,2514	0,1885
X30	0,0000	0,6655	0,0669	0,0240	0,1095
X31	0,0000	0,8038	0,1273	0,0555	0,1705
X32	0,0000	1,0000	0,6999	0,7677	0,2454

TABLE 3. Table over min, max, mean, median and standard deviation for all variables.

4.2 ASSUMPTION REQUIREMENTS

In all the following tests that are being presented, a 95% significance level is used. In table 4, the probability value for t-test, Mann Whitney U test, White's heteroskedasticity test and Jarque-Bera are summarized.

The t-test was applied to examine which variables are statistical significant. The null hypothesis is the following:

H_0: No significance difference between the two groups.

The null hypothesis can not be rejected for the variables X13, X15, X16, X18, X20, X23, X24, X25, X26 X28, X29, X30 and X31.

One more test that is necessary to be applied in order to be able to use the Multiple Discriminant Analysis method (MDA) is the Mann Whitney U test. According to the SPSS manual, this test is non parametric and is used to examine if two samples of observations belong to the same distribution. The null hypothesis that is used in the Mann Whittney U test is the following:

H_0: The two samples of observations come from the same population.

The results that were obtained allow us to reject the null hypothesis for the variables X17, X20, X23, X28, X29, X30, X31 and X32.

Another test employed is White's test for heteroskedasticity. The purpose is to examine if the errors have constant variance and if they do not then there are evidence of heteroskedasticity. Homoskedasticity is required when Multiple Discriminant Analysis (MDA) and logit are used to predict firm bankruptcy. The null hypothesis is the following (Eviews manual):

H_0: No indication of heteroskedasticity (homoskedasticity)

The results that were obtained indicate that the variable X30 and X31 has heteroskedasticity evidence. In other word, the null hypothesis is rejected for the variables X30 and X31

Finally, the Jarque-Bera test was applied in order to examine if the data sample is normally distributed. Normal distribution assumption is necessary if Multiple Discriminant Analysis (MDA) is applied. The null hypothesis is the following:

H_0: The data sample is normally distributed and the Jarque-Bera statistic is distributed as χ^2 with 2 degrees of freedom.

The results that are obtained allow us to reject the null hypothesis for all the variables and thus the variables are not normally distributed. The software that is used in this study is thus used to relax this assumptions when applying the empirical study.

TABLE 4

ASSUMPTION TESTING FOR THE VARIABLES

Variable	*t-test*	*MW U-test**	*White***	*Jarque-Bera****
X1	0,003	0,044	0,5986	0,000
X2	0,001	0,000	0,0747	0,000
X3	0,000	0,000	0.0844	0,000
X4	0,001	0,000	0.8446	0,000
X5	0,000	0,000	0.8175	0,000
X6	0,006	0,000	0.9374	0,000
X7	0,000	0,000	0.3009	0,000
X8	0,003	0,000	0.1744	0,000
X9	0,004	0,000	0.1203	0,000
X10	0,003	0,000	0.5565	0,000
X11	0,000	0,000	0.1506	0,000
X12	0,012	0,000	0.0527	0,000
X13	0,069	0,044	0.4194	0,000
X14	0,008	0,000	0.8744	0,000
X15	0,394	0,000	0.5884	0,000
X16	0,348	0,000	0.9006	0,000
X17	0,019	0,056	0.9313	0,000
X18	0,435	0,004	0.9739	0,000
X19	0,000	0,000	0.4256	$1,48*10^{-5}$
X20	0,541	0,919	0.7459	0,000
X21	0,005	0,012	0.6017	0,000
X22	0,000	0,000	0.9938	0,000
X23	0,932	0,176	0.2529	0,000
X24	0,190	0,000	0.4084	0,000
X25	0,051	0,002	0.3265	0,000
X26	0,120	0,000	0.5437	0,000
X27	0,006	0,004	0.1429	0,000
X28	0,877	0,397	0.5262	0,000
X29	0,566	0,623	0.8569	0,004
X30	0,917	0,960	0.0003	0,000
X31	0,793	0,942	0.0333	0,000
X32	0,000	0,891	0.3769	0.0001

* Mann-Whitney U-test; Non-parametrical test for group separation
** White's test for heteroskedasticity
*** Jarque-Bera test for normality

TABLE 4. Various test-statistics for the variables

4.3 MDA RESULTS

(i) Model performance

In assessing the model and evaluating the results the discriminant function is applied to out-of-sample firms and the model development is obtained from the in-sample firms. The final discriminant function is given by:

$$Z_{swe} = -0{,}421X_2 + 4{,}785X_5 - 0{,}461X_{19} + 5{,}550X_{22} + 0{,}280X_{30}$$

where

X_2 = EBITA to Total Assets
X_5 = Equity Share of Total Assets (Solvency)
X_{19} = Firm Size (Natural Logarithm of Sales)
X_{22} = Short and Long Term Debt to Total Assets
X_{30} = Inventory to Sales
Z_{swe} = Swedish Z-score

The cut-off score is calculated at 0,323 and indicates that firm discriminant scores over the critical score are categorized as defaulted and firm discriminant scores under the cut-off point are categorized as non-defaulted. Firms that actually defaulted but is according to the MDA model categorized as non-default falls under the type I error category, and firms that are healthy but according to the MDA-model classed as defaulted falls under the type II error category. The following table shows type I and type II errors of the in-sample firms:

	Predicted		
Actual	Group I	Group II	Total
Group I	72	18	90
Group II	23	67	90

The model accurately categorizes defaulted firms to 80,0% while the accuracy for non-defaulted firms is less at 74,4%. The results of the multiple discriminant analysis point to difficulties in correctly categorizing non-defaulted firms correctly. The in-sample firms are accurately categorized in 77,2% of the cases for all the 180 firms used for model development. That is roughly 3 out of 4 firms can accurately be categorized as either defaulted or non-defaulted.

Table 5 shows the distribution of discriminant scores for the in-sample firms in different classes, with the cut-off at 0,323:

TABLE 5

DISCRIMINANT SCORES OVER IN-SAMPLE FIRMS

Z-score	*Defaulted*	*Non-defaulted*
(2,60) – (2,00)	0,00%	1,11%
(2,00) – (1,00)	0,00%	2,22%
(1,00) – (0,00)	11,11%	52,22%
0,00 – **0,323**	8,89%	18,89%
0,323 – 0,50	8,89%	10,00%
0,50 – 1,00	48,89%	13,33%
1,00 – 1,50	14,44%	2,22%
1,50 – 2,00	4,44%	0,00%
2,00 – 3,20	3,33%	0,00%
	100,00%	**100,00%**

TABLE 5. The z-score range and the categorizations of the in-sample firms with the cut-off at 0,323.

The discriminant scores for the in-sample firms are concentrated around the cut-off point at 0,323 and 86,1% off all firms have z-scored between -1,00 to 1,00. Scores in the nearest vicinity of the cut-off score could be considered "grey area" scores, meaning that they are associated with a high level of uncertainty even though accurately categorized. Considering the cut-off score at 0,323 the grey area is between 0 and 0,5 which in case 23,3% of the firms fall into the uncertainty category.

(i) Model holdout testing

Probably the most important step in any model building for prediction of classification is the model testing. Testing the model on a holdout sample not included in the actual model development gives a clear picture of the model robustness. The following table specifies type I and type II errors among the holdout firms when applying the above stated discriminant score to them:

	Predicted		
Actual	Group I	Group II	Total
Group I	18	5	23
Group II	7	16	23

The model shows hardship in holdout accuracy and the overall accuracy in prediction capability remains lower than that for the actual in-sample model performance at 73,9% compared 77,2%. Yet again the model is better in accurately categorizing defaulted firms rather than non-defaulted firms. The accuracy of bankruptcy categorization is 78,3% compared to 69,6% for the healthy firm categorization. Clearly the model suffers from predictive ability even though the accuracy is far beyond 50%.

TABLE 6

DISCRIMINANT SCORES OVER HOLDOUT-SAMPLE FIRMS

Z-score	*Defaulted*	*Non-defaulted*
(2,60) – (2,00)	0,00%	0,00%
(2,00) – (1,00)	0,00%	0,00%
(1,00) – (0,00)	8,70%	39,13%
0,00 – **0,323**	13,04%	30,43%
0,323 – 0,50	4,35%	17,39%
0,50 – 1,00	39,13%	8,70%
1,00 – 1,50	26,09%	4,35%
1,50 – 2,00	8,70%	0,00%
2,00 – 3,20	0,00%	0,00%
	100,00%	**100,00%**

TABLE 6. The z-score spans and the categorizations of the holdout-sample firms with the cut-off at 0,323.

The distribution of the z-scores among the holdout-sample firms is also concentrated around the cut-off score with very few observations of extreme scores in either direction. However the model has 26,1% of the defaulted firms in the 1,00-1,50 score category compared to 14,4% the in-sample firms that fall in the same category. This point to a greater ability to accurately predict firms further away from the grey area than the actual model for in-sample firms does. As a result of this, for the overall model 80,4% of all firms falls in the -1,00 to 1,00 z-score category as compared to 86,1% of the in-sample firms that falls into this category. Further, as much as 95,7% of the non-defaulted firms fall into this category. The grey area firms amounts to 32,6% of the total holdout-sample compared to 23,3% of the in-sample firms, thus the uncertainty of the holdout model must be considered greater. However only 17,4% of the defaulted firms falls into the grey area category while 47,8% of the non-defaulted firms falls into this category. Once again the model points to a higher accuracy in accurately categorizing defaulted firms into its correct category. In other words the frequency of type II errors is higher than that for type I errors.

4.4 LOGIT RESULTS

4.4.1 STANDARD MODEL

(i) Model performance
The standard model is developed by only including only variables that are significant on at least the 90% level. No model including variables at least significant on the 80% level is run since no variables with a significance level between 80% and 90% are found. The following logit function is obtained:

$$\ln\left[\frac{P}{1-P}\right] = 7{,}972 + 0{,}227X_1 - 5{,}905X_3 - 7{,}685X_7 + 0{,}0025X_{14} + 0{,}0287X_{16} - 0{,}720X_{19}$$

where

X_1 = Sales to Total Assets
X_3 = Retained Earnings to Total Assets
X_7 = Return on Assets (ROA)
X_{14} = EBITA per Employee
X_{16} = Change in Total Assets from Last Year
X_{19} = Firm Size (Natural Logarithm of Sales)

The cut-off score is naturally given by 0,5 thus a logit score of over 0,5 is categorized as a defaulted firm and a logit score under 0,5 is considered a healthy firm. The following shows a table over in-sample type I and type II errors in terms of the number of firms for the logit model in question:

	Predicted		
Actual	Group I	Group II	Total
Group I	73	17	90
Group II	16	74	90

Overall accuracy of the model is thus 81,67%. Roughly 8 out of 10 firms are categorized accurately both for bankruptcy cases and healthy firms. There is no basic difference between the ability to categorize either firm group more accurately. The type I error frequency is 18,9% which is slightly higher than for the type II errors of 17,8%.

Table 7 shows the logit score classification of the in-sample firms:

TABLE 7

LOGISTIC SCORES OVER IN-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	1,11%	26,67%
0,10 – 0,20	2,22%	28,89%
0,20 – 0,30	2,22%	8,89%
0,30 – 0,40	5,56%	11,11%
0,40 – **0,50**	7,78%	6,67%
0,50 – 0,60	11,11%	8,89%
0,60 – 0,70	13,33%	2,22%
0,70 – 0,80	12,22%	2,22%
0,80 – 0,90	7,78%	1,11%
0,90 – 1,00	36,67%	3,33%
	100,00%	**100,00%**

TABLE 7. Score distribution for the standard logistic regression of in-sample firms.

The distribution of logit scores has a clear and logical tendency toward extreme values both in the case of defaulted firms and non-defaulted firms. The 20th percentile (0 – 0,2 and 0,8 – 1,0) in both directions contain 50% of the total in-sample firms. Further the 10th percentile in both directions contain 31,7% of the total in-sample firms. As can be seen in the table above the model shows a pattern of increasing frequencies of firms toward the boundary values of 1 in the bankruptcy cases and 0 in the cases of healthy firms. The grey area in the logit case is defined as 0,4 – 0,6 probability of bankruptcy. Of the total sample 17,2% of the firms fall into the uncertain grey area group.

(ii) Model holdout testing

Logit model testing is carried out by applying the logit probability function to the holdout-sample firms. The following type I and type II errors is obtained:

	Predicted		
Actual	Group I	Group II	Total
Group I	20	3	23
Group II	3	20	23

The model categorizes similarly for both groups in the holdout sample and the overall predictive accuracy is 87% and the error rate for both type I and type II is 13%. Thus the predictive accuracy is higher than the in-sample categorization accuracy of 81,7% which is a striking and unanticipated result.

The classification table yields the following results:

TABLE 8

LOGISTIC SCORES OVER HOLDOUT-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	4,35%	17,39%
0,10 – 0,20	4,35%	21,74%
0,20 – 0,30	0,00%	4,35%
0,30 – 0,40	0,00%	17,39%
0,40 – **0,50**	4,35%	26,09%
0,50 – 0,60	17,39%	0,00%
0,60 – 0,70	8,70%	0,00%
0,70 – 0,80	8,70%	4,35%
0,80 – 0,90	8,70%	4,35%
0,90 – 1,00	43,48%	4,35%
	100,00%	**100,00%**

TABLE 8. Score distribution for the standard logistic regression of holdout-sample firms.

The holdout sample scores show even a clearer bias toward extreme values in the bankruptcy firm case. The 10th percentile contain 43,48% and 17,39% respectively for the defaulted and non-defaulted firms. For the total sample the percentage of firms in the 10th percentile in both directions is 30,4%, thus much like the in-sample firms. The total percentage of holdout firms in the 20th both directions amounts to 45,7% which is slightly lower than for the in-sample firms. A total of 11 of 46 firms, which corresponds to 23,9%, are found having a grey area score between 0,4 and 0,6. Thus the holdout sample displays higher uncertainty than the in-sample.

4.4.2 FORWARD CONDITIONAL LOGISTIC MODEL

(i) Model performance

The forward conditional logistic regression includes variables in a stepwise manor, so that the following logistic function is obtained:

$$\ln\left[\frac{P}{1-P}\right] = 10{,}302 - 4{,}662X_2 - 8{,}624X_5 + 1{,}896X_9 - 0{,}768X_{19} - 0{,}017X_{23} + 4{,}184X_{30}$$

where

X_2 = EBITA to Total Assets
X_5 = Equity Share of Total Assets (Solvency)
X_9 = Gross Profit Margin
X_{19} = Firm Size (Natural Logarithm of Sales)

X_{23} = Short and Long Term Debt to Equity
X_{30} = Inventory to Sales

Much like the previous logistic function this also has six explanatory variables, however only one variable in common and that is firms size measured as the natural logarithm of sales. The following classification accuracy is obtained for the forward conditional logit model:

	Predicted		
Actual	Group I	Group II	Total
Group I	72	18	90
Group II	17	73	90

The model gives an overall accuracy of 80,6% for correct classification. For the bankruptcy group the accuracy is 80,0% and for the healthy firms the number is 81,1%, thus the model is slightly better at categorizing non-defaulted firms than defaulted firms.

The complete classification table follows:

TABLE 9

LOGISTIC SCORES OVER IN-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	1,11%	36,67%
0,10 – 0,20	3,33%	14,44%
0,20 – 0,30	4,44%	13,33%
0,30 – 0,40	4,44%	5,56%
0,40 – **0,50**	6,67%	11,11%
0,50 – 0,60	4,44%	6,67%
0,60 – 0,70	6,67%	5,56%
0,70 – 0,80	22,22%	4,44%
0,80 – 0,90	13,33%	0,00%
0,90 – 1,00	33,33%	2,22%
	100,00%	**100,00%**

TABLE 9. Score distribution for the forward conditional logistic regression of in-sample firms.

Logically the results show a clear bias toward defaulted firm logit scores toward 1 and 0 for the non-defaulted. For both groups the 10th percentile contains 35% of the firm which is slightly higher than for the previous logit model. For the 20th percentile this number is 48,9% which is also higher than the previous model. Firms that fall into the uncertainty category amounts to 26 observations, which is 14,4% of the total sample.

(ii) Model holdout testing
The above stated model for the forward conditional logistic regression is applied to the 46 holdout sample firms. The error distribution is as follows:

	Predicted		
Actual	Group I	Group II	Total
Group I	16	7	23
Group II	6	17	23

For the forward conditional model the loss of predictive power is apparent. From an in-sample accuracy of 80,6% it has dropped roughly 10 points to 71,7% for the holdout-sample. The model testing shows that the model is slightly better in categorizing non-defaulted firms than defaulted firms.

TABLE 10

LOGISTIC SCORES OVER HOLDOUT-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	4,35%	30,43%
0,10 – 0,20	13,04%	21,74%
0,20 – 0,30	4,35%	0,00%
0,30 – 0,40	8,70%	13,04%
0,40 – **0,50**	0,00%	8,70%
0,50 – 0,60	0,00%	8,70%
0,60 – 0,70	4,35%	8,70%
0,70 – 0,80	13,04%	0,00%
0,80 – 0,90	17,39%	4,35%
0,90 – 1,00	34,78%	4,35%
	100,00%	**100,00%**

TABLE 10. Score distribution for the forward conditional logistic regression of holdout-sample firms.

For the holdout sample the logit scores are more dispersed than for the previous model. Even though the model testing shows little predictive ability for the forward conditional compared to previous models, one of the strengths of it seem to be the ability of not categorizing defaulted firms in the uncertainty range. The results show that 32,6% of the firms fall into the 10th percentile of corner values, 1 and 0, for all the firms and 52,2% for the 20th percentile. Both of these numbers are fairly high compared to previous models, which is somewhat contradictory to the fact that the model has lower predictive in comparison. The number of firms that is found in the uncertainty area amount to only 4 firms, which corresponds to a percentage of 8,7% for the total sample.

4.4.3 BACKWARD CONDITIONAL LOGISTIC MODEL

(i) Model performance

The backward conditional logistic regression omits variables in a stepwise manor, so that the following logistic function is obtained:

$$\ln\left[\frac{P}{1-P}\right] = 11{,}234 - 0{,}030X_4 - 6{,}872X_5 - 12{,}528X_7 + 5{,}993X_9 + 0{,}002X_{14} - 0{,}824X_{19} - 0{,}020X_{23} + 3{,}980X_{30}$$

where

X_4 = Firm Age
X_5 = Equity Share of Total Assets (Solvency)
X_7 = Return on Assets (ROA)
X_9 = Gross Profit Margin
X_{14} = EBITA per Employee
X_{19} = Firm Size (Natural Logarithm of Sales)
X_{23} = Short and Long Term Debt to Equity
X_{30} = Inventory to Sales

This model employs more explanatory variables than the previous models have done. The type I and type II error frequency of the model is listed below:

	Predicted		
Actual	Group I	Group II	Total
Group I	76	14	90
Group II	11	79	90

Overall accuracy for the backward conditional model is 86,1%, which is a relative high in-sample accuracy. The model seem to be better in classifying non-defaulted firms correctly as they are accurately categorized in 87,8% of the cases or roughly 9 firms out of 10. Accuracy for the defaulted firms remains lower at 84,4%.

All in all the in-sample categorization is high and in Table X follows the complete distribution over the logit score frequencies for the backward conditional model:

TABLE 11

LOGISTIC SCORES OVER IN-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	2,22%	46,67%
0,10 – 0,20	2,22%	10,00%
0,20 – 0,30	2,22%	11,11%
0,30 – 0,40	1,11%	10,00%
0,40 – **0,50**	7,78%	10,00%
0,50 – 0,60	6,67%	1,11%
0,60 – 0,70	6,67%	6,67%
0,70 – 0,80	18,89%	2,22%
0,80 – 0,90	8,89%	0,00%
0,90 – 1,00	43,33%	2,22%
	100,00%	**100,00%**

TABLE 11. Score distribution for the backward conditional logistic regression of in-sample firms.

The distribution of logit scores for the backward conditional model displays clear bias toward the extreme values and 45% of the total firms logit scores can be found in the 10th percentile. This is relatively a high number compared to the other models and shows an apparent robustness in in-sample categorization. Looking at the 20th percentile the firm score distribution takes on an more modest frequency at 54,4%. This fact however further understates the models clear bias of accurately categorizing in-sample firms. Firms that fall into the uncertainty category are also fewer than average of the previous logistic models and 23 firms or 12,8% fall into this category.

(ii) Model holdout testing

The model holdout testing yielded the following type I and type II error frequencies:

	Predicted		
Actual	Group I	Group II	Total
Group I	17	6	23
Group II	5	18	23

Test results show the models obvious loss in predictive capability. From being a highly accurate in-sample categorization model the prediction accuracy is only 76,1%, which is a drop of 10 points from the in-sample accuracy of 86,1%. The model remains better in categorizing of non-defaulted firms into the right category compared to defaulted firms, however the difference is minimal.

In Table 12 the complete distribution of holdout firms scores are displayed:

TABLE 12

LOGISTIC SCORES OVER HOLDOUT-SAMPLE FIRMS

Logit-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	13,04%	34,78%
0,10 – 0,20	4,35%	8,70%
0,20 – 0,30	4,35%	13,04%
0,30 – 0,40	0,00%	8,70%
0,40 – **0,50**	4,35%	13,04%
0,50 – 0,60	4,35%	4,35%
0,60 – 0,70	0,00%	8,70%
0,70 – 0,80	17,39%	0,00%
0,80 – 0,90	17,39%	0,00%
0,90 – 1,00	34,78%	8,70%
	100,00%	**100,00%**

TABLE 12. Score distribution for the backward conditional logistic regression of holdout-sample firms.

For the holdout sample the distribution of scores are somewhat dispersed and no clear logical pattern can be discerned in comparison to previous models. In the 10th percentile a total of 16 firms is found, which is equivalent to 34,8%. In the 20th percentile the total number of accurately categorized firms are 22 or 47,8%. The number of firms that fall into the uncertainty category amounts to 6 or 13,04%.

4.5 NEURAL NETWORKS RESULTS

(i) Model performance

Feedforward, feedback and probabilistic are the three different types of neural networks that were used to predict corporate bankruptcy. The networks are programmed in a binary fashion so that the defaulted firm are encoded as 1 and non-defaulted as 0, naturally it follows that the cut off point is 0,5 analogous to the logistic regression.

The type I and type II error frequencies for the observations of the in the sample are the following:

	Predicted		
Actual	Group I	Group II	Total
Group I	65	25	90
Group II	11	79	90

All the three models have the same type I and type II error at in the sample. The type I error is 27,77% (2.22%+2.22%+4.44%+7.78%+11.11%) and the type II error is 12,22% (6.67%+2.22%+2.22%+0.00%+1.11%).

The scores that the methods obtained for the in-sample can be seen in Appendix IV. We have summarized the classification results that were obtained for the in-sample firms in Table 13. All three methods obtain the same classification results at in the sample. A non-defaulted firm belongs to the range 0 to 0,5 and a defaulted belongs to 0,5-1. At in the sample, the models classified correctly 78,26% of the defaulted firms and 87,78% of the non-defaulted firms.

TABLE 13

NEURAL NETWORK SCORES OVER IN-SAMPLE FIRMS

NN-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	2.22%	36.67%
0,10 – 0,20	2.22%	20.00%
0,20 – 0,30	4.44%	14.44%
0,30 – 0,40	7.78%	6.67%
0,40 – **0,50**	11.11%	10.00%
0,50 – 0,60	14.44%	6.67%
0,60 – 0,70	16.67%	2.22%
0,70 – 0,80	13.33%	2.22%
0,80 – 0,90	10.00%	0,00%
0,90 – 1,00	17.78%	1.11%
	100,00%	**100,00%**

TABLE 13. Score distribution for all the neural networks of in-sample firms.

(ii) Model holdout testing

The type I and type II error frequencies for the observations of the out of the sample are the following:

	Predicted		
Actual	Group I	Group II	Total
Group I	18	5	23
Group II	4	19	23

The type I error is 21,75% (0.00%+4.35%+4.35%+4.35%+8.70%) and the type II error is 17,39% (13.04%+0.00%+0.00%+4.35%+0.00%).

The scores that the three methods obtain for the holdout-sample can be seen in Appendix IV. We have summarized the classification results that were obtained for the out of the sample in Table 14. All three methods obtained the same classification results for the out-of-sample as well. At the out of the sample the models classified correctly 78,26% of the defaulted firms and 82,62% of the non-defaulted firms.

TABLE 14

NEURAL NETWORK SCORES OVER HOLDOUT-SAMPLE FIRMS

NN-score	*Defaulted*	*Non-defaulted*
0,00 – 0,10	0.00%	26.09%
0,10 – 0,20	4,35%	4.35%
0,20 – 0,30	4,35%	26.09%
0,30 – 0,40	4.35%	8,70%
0,40 – **0,50**	8.70%	17.39%
0,50 – 0,60	8.70%	13.04%
0,60 – 0,70	13.04%	0.00%
0,70 – 0,80	17,39%	0,00%
0,80 – 0,90	17,39%	4.35%
0,90 – 1,00	21.74%	0.00%
	100,00%	**100,00%**

TABLE 14. Score distribution for the forward conditional logistic regression of holdout-sample firms.

4.6 SUMMARY OF RESULTS

TABLE 15

SUMMARY OF RESULTS FOR THE MODELS

IN-SAMPLE

Model	*Overall Accuracy*	*Type I Error*	*Type II Error*	*% in Grey Area*
Discriminant Analysis	77,22%	20,00%	25,56%	23,34%
Logistic Regression				
- Standard	81,67%	18,89%	17,77%	17,23%
- Forward Conditional	80,55%	20,00%	18,89%	14,44%
- Backward Conditional	86,12%	15,55%	12,22%	12,78%
Neural Network				
- Feedforward	80,00%	27,77%	12,22%	21,11%
- Backward	80,00%	27,77%	12,22%	21,11%
- Probabilistic	80,00%	27,77%	12,22%	21,11%

HOLDOUT SAMPLE

Model	*Overall Accuracy*	*Type I Error*	*Type II Error*	*% in Grey Area*
Discriminant Analysis	73,91%	21,74%	30,44%	32,61%
Logistic Regression				
- Standard	86,97%	13,04%	13,04%	23,92%
- Forward Conditional	71,74%	30,44%	26,10%	8,70%
- Backward Conditional	76,09%	26,09%	21,75%	13,04%
Neural Network				
- Feedforward	80,44%	21,75%	17,39%	23,92%
- Backward	80,44%	21,75%	17,39%	23,92%
- Probabilistic	80,44%	21,75%	17,39%	23,92%

TABLE 15. An overview of the obtained results in this study.

CHAPTER 5
ANALYSIS

The analysis chapter gives a short re-cap of the results and sets out to explain the obtained results. Results that are produced in this study are compared to results obtained in prior research. Further explanations on why the models produces the given results are presented as well as possible drawbacks and advantages with the various models.

The in-sample for all three versions of logistic regression is better in categorizing firm bankruptcy than the three versions of neural network (feedforward, feedback and probabilistic) and multiple dicriminant analysis (MDA). But it is interesting to observe that for the holdout-sample the forward and backward conditional logistic regression have a deterioration in prediction accuracy of 8,81% and 10,03% respectively. Only the standard logistic regression's prediction accuracy increased and the improvement that was observed is 5,3%. The neural networks slightly increased for the out-of-sample and the multiple discriminant analysis (MDA) is reduced by 3,31%. The forward conditional logistic regression which was the third best model in categorizing bankruptcy, it was the second worst in predicting bankruptcy.

It is also of interest to examine the type I and type II error frequency for each model because of the serious consequences that they might cause lenders. Type I error is detected when the applied model is classifying a bankrupt firm as a non-bankrupt firm and in the case of type II error the model is classifying a non-bankrupted firm as a bankrupt. If a lender is applying a model with high type I error frequency it will probably turn out to be costly because according to the model the firm will be able to repay its debts but in reality the firm will go bankrupt. On the other hand, if a lender is using a model with high type II error that will reduce the lender's profits because the model will conclude that the firm is not trustworthy but in reality this is not true. The standard logistic regression is the only model which has a prediction accuracy that increases and at the same time both type I and type II errors improve. At in the sample it has a type I error of 18,89% and it decreases to 13,04%. All the other models have a type I error that is above 20%. Moreover, the in-sample has a type II error of 17,77% that decreases to 13,04% which is a satisfying result compared to the type II errors that the other models obtain. The backward conditional logistic regression model is the best model for overall in-sample accuracy, type I and type II errors. This model has the lowest type I and type II error for the in-sample but the out-of-sample it has the second highest type I error (it increased by 10,54%) and the third highest type II error (it increased by 9,53%). Even though the results are satisfying for the in-sample group, for the holdout-sample it turned out to be one of the worst models for a lender to apply. Furthermore, it is not only the prediction accuracy of the forward conditional logistic regression that decreases but also the type I and type II errors deteriorate at out of the sample. The type I error decreases with 6,96% and type II error with 7,21%. This model also turned out to be one of the worst models that a lender could apply. The type I error at the discriminant analysis method increased only with 1,74% but it was already high at the in the sample (20%). The discriminant analysis method had already the highest type II error at the in the sample and it deteriorated even more at the out of the sample, it increased to 30,44%. As it can be observed the results are not encouraging.

Finally, all three neural networks obtain the same type I and type II errors both at in the sample and out of the sample. Those models have better type I error in the out of the sample compared to the in the sample but type II error increased with 5,17%. The three neural networks are the second best models that a lender could use in order to predict corporate bankruptcy.

Before carrying out any further analysis of the models and comparing the results to prior research it is important to show the relevancy of the obtained results. This is done by examining the variable signs in the various functional form models (MDA and Logit). The models could potentially have very high categorization accuracy but lack explanatory power due to the irrelevancy that the models display. In Table 16 below a summary of the included variables in the functional form models are presented along with their expected and obtained signs. This is followed by a stepwise analysis of the relevancy of each variable, and a possible explanation of why we obtain unexpected results. This relevancy analysis is important when further assessing and analyzing the models performance and accuracy.

TABLE 16

VARIABLE SPECIFICATION OF MDA AND LOGIT

Model	*Variables*	*Expected Sign*	*Obtained sign*	*Inference*
Discriminant Analysis	X2	-	-	Expected
	X5	-	+	Unexpected
	X19	-	-	Expected
	X22	+	+	Expected
	X30	+	+	Expected
Logistic Regression				
Standard	X1	-	+	Unexpected
	X3	-	-	Expected
	X7	-	-	Expected
	X14	-	+	Unexpected
	X16	+/-	+	N/A
	X19	-	-	Expected
Forward Conditional	X2	-	-	Expected
	X5	-	-	Expected
	X9	-	+	Unexpected
	X19	-	-	Expected
	X23	+	-	Unexpected
	X30	+	+	Expected
Backward Conditional	X4	-	-	Expected
	X5	-	-	Expected
	X7	-	-	Expected
	X9	-	+	Unexpected
	X14	-	+	Unexpected
	X19	-	-	Expected
	X23	+	-	Unexpected
	X30	+	+	Expected

TABLE 16. Expected sign and obtained sign for the variables included in the final MDA and Logit models.

X1 – Sales to Total Assets. A large portion of sales in relation to the asset base means a large operational efficiency and an ability to generate sales on the firm's current asset base. This capital turnover ratio is expected to be lower for defaulting firms thus yielding a negative sign. However the results show the opposite which is unexpected. It is not entirely easy to explain this phenomenon. One explanation could be that a defaulting firm is forced to sell of a lot of its balance sheet assets in order to satisfy the creditors demand for interest payments, while maintaining a stable turnover, thus actually increasing the ratio.

X2 – EBITA to Total Assets. The ratio indicates how capable the firm is in generating earnings in relation to their asset base. The obtained sign indicates that the more earnings in relation to the total assets decrease the bankruptcy propensity and this is an expected inference.

X3 – Retained Earnings to Total Assets. This ratio expresses the amount of accumulated earnings in relation to the assets that the firm has generated during the course of its lifetime. It is expected that the more earnings that a firm has as reserve the more likely the firm is to sustain an economic downturn, hence the more retained earnings a firm has to its asset base the less likely it is to default. The obtained negative sign is also expected.

X4 – Firm Age. The firm age displays the number of years that the firm has been in operations. Logically, a firm operating in the beginning of its lifecycle is more likely to go bankrupt due to for instance the necessity of capital input during this period. The variable has a negative sign which is also expected.

X5 – Equity Share of Total Assets (Solvency). Solvency is a classic leverage ratio which displays the firm financing structure. It is expected that the more equity financed a firm is the less likely it is to default, due to the lower outlay of continuous liability payments. The obtained sign displays a contradictory inference and is unexpected. This fact can be explain the complementary portion of financing through liabilities to be interest free, yielding a large portion of short term, interest-free, financing.

X7 – Return on Assets (ROA). This is a commonly used profitability measure. In general, for firms with a higher propensity of defaulting, the profitability should be lower thus a negative sign applies. The obtained sign is negative thus consistent with the hypothesis.

X9 – Gross Profit Margin. The gross profit margin indicates how capable a firm is to generate profits on its sales. It is expected that defaulting firms has greater hardships to generate profits in relation to its sales due to a higher cost base. Our results yield a positive coefficient for this variable, thus contradictory to what would be expected. The only reasonable explanation for this is the fact that gross profit margin is calculated as an EBITA to Sales ratio, which in fact excludes possible liability payments. Defaulted firms can still be as profitable (or more) than healthy firms on a gross-profit basis, however if including financial expenses the result might be different.

X14 – EBITA per Employee. This ratio measures how much earnings are generated in relation to the number of employees the firm has. This form of efficiency measure should be lower for defaulting firms, thus yielding a negative sign. However the obtained results yield a negative sign, which is unexpected. A possible explanation could be, yet again, that the EBITA measures profit before financial items, which in the case of a defaulting firm would be greater than for a healthy firm.

X16 – Change in Total Assets from Last Year. Measures a change in the firm's asset base from one year to another. A positive change of this variable could mean one of two things: (i) The firm increases its assets base on the basis of profitable investments or; (ii) The firm increases its asset base but have difficulties in profitably deploying the growing asset base. Thus the variable could take on both a negative and positive sign for either of the two firm groups. Here, the variable is of positive nature which indicates the latter of above stated cases.

X19 – Firm Size (Natural Logarithm of Sales). The firm size is an important firm characteristic which in this case is measured by its sales. It is anticipated that larger companies have greater resources to sustain and battle a bankruptcy situation, thus the variable is expected to take on a negative sign. The hypothesis is consistent with the obtained results.

X22 – Short & Long Term Debt to Total Assets. This ratio is a measure of the firm financing structure. Carrying a high fraction of liabilities on the total asset base, would logically mean higher interest payments, which in turn increases the probability of bankruptcy. The results support this hypothesis as the variable coefficient has a positive sign.

X23 – Short & Long Term Debt to Equity. This leverage ratio indicates the relation between liabilities and equity capital used in financing the firms operations. Increasing the leverage ratio would infer an increased likelihood of bankruptcy, thus a positive sign for the variable coefficient. However, the obtained results display the contrary which is unexpected. The explanation for this could be that, as mentioned above, the defaulting firm has the majority of its liabilities tied to interest-free short term debt (such as accounts payable).

X30 – Inventory to Sales. This ratio measures the management's ability to run the firm's operations efficiently with as little binding of capital as possible (such as storage and supplies) in relation to how much the firm turnover. It is expected that defaulting firms are less liquid than healthy firms thus less binding of capital. The hypothesis is therefore that the sign should be positive which the results also show.

Studying the MDA and the logit models some variables are recurrent throughout the various model types. The most persistent variable is X_{19} (firm size measured as the natural logarithm of sales) which is present in both the MDA and all of the logistic regression types. The firm size coefficient is also consistently negative, thus indicating that smaller firms have higher propensity of bankruptcy. There is a high relative high level of unexpected signs in the resulting models. The unexpected sign frequency is 32% and must be considered when applying the models in real-world situations. Otherwise, the model interpretation and relevance might be overrated.

A loss in predictive power for all the models should be expected, however for both the standard logistic model and the neural networks the categorization accuracy increase for the holdout sample firms. Although the neural networks only increase marginally the results are prevalent; Logistic regression and neural networks clearly dominate multiple discriminant analysis as a assessment tool for predicting corporate bankruptcy. The relative poor performance of the MDA could however partially be explained by the inclusion of variable X_{30} (Inventory to Sales) which shows evidence of heteroskedasticity being present. Further, all the variables need to have a normal multivariate distribution (Serrano-Cinca 1997) and the data sample in this thesis is not normally distributed. This is one of the main reasons why the discriminant analysis fails to predict corporate bankruptcy so well as the other models. Determining the superiority between logistic regression and neural networks is not as obvious. On the one hand the logit models have higher in-sample accuracy, but lose significant predictive power applying them to the holdout sample (except for the standard logistic model), whereas the neural networks categorically improve its holdout-sample accuracy. Also the neural network models have a lower type I error categorization than the two latter logit models. A type I error is the most costly error from a credit grantor point of view since this error categorizes unsound firms into the healthy firm category, thus substantial sums could

potentially be lost. A robust trustworthy bankruptcy prediction model should preferably have a higher larger portion of type II errors in relation to the total errors. The single best model for predicting bankruptcy is, however, the standard logistic regression only including highly significant variables. It gives a high accuracy both for in-sample firms and holdout sample firms. The high predictive accuracy of 87% for this model shows a clear robustness in bankruptcy prediction. In prior research the general conclusion is MDA<Logit<ANN in terms of model ranking conditional on the model predictive accuracy (Altman et al 1994; Back et al 1997; Etheridge et al 1997). Thus the fact that a logistic regression model has the best predictive accuracy is somewhat unexpected. Logistic regression has potentially a larger grey area than for instance neural network, due to its sigmoid function form. This infers that logistic regression "needs" more of its observations in the tails, thus a more extreme value bias, than neural networks. A cut-off point of 0,5 might be too narrow as a critical score since logistic regression has a large linear area (see Figure 4 for deeper understanding) trying to capture non-linear data. This could be one explanation why our results differ from prior research findings showing neural networks being the superior model. However, the standard logistic model is the most extreme value biased of all the functional form models, thus shows evidence of high categorization ability compared to neural networks. The various neural networks yield the same results both with regards of in-sample categorization and holdout-sample categorization. This show strength and reliability among the neural network models in terms of robustness and predictability. In explaining why the different NN-models yield the same results, the amount of data is essential. In a large data set (hundreds of thousand or even millions of data points) the neural network is able to track more learning patterns than for smaller data-sets. Apparently, the data-set in our study is restricted to 5760 in-sample data-points (used to train the network) thus the various NN-models track the exact same patterns or characterizing data paths. The restricted amount of data could also potentially explain why the neural network is underperforming in relation to the standard logistic regression. The logistic regression and the discriminant analysis do not require a large amount of observations in order to perform well as the neural network models require (Kutner et al 2004). The logistic regression and the discriminant analysis are models which were developed for data samples which do not contain more than 1000 observations (Kutner et al 2004). Probably, the neural network models did not have enough of observations (input) so that they could identify enough of patterns between the two category groups. Moreover, a more thorough evaluation and trial-and-error testing can be implemented with regards to number of neurons in the hidden layer and optimal learning time As a result the neural network models could not train and test enough in order to be able to predict corporate bankruptcy more accurately as some prior research indicates.

It can generally be said that the findings in this study largely correspond to findings in prior research both in model ranking and in individual model accuracy. The range for accurate model prediction in this study stretches from 74% to 87%. From this it is concluded that generally quantitative bankruptcy models are a useful tools in assessing corporate bankruptcy probabilities. The results show upon a higher accuracy classification than chance which is satisfactory, however exclusively using the tools presented in this study for categorizing good and bad loan applicants conditional on their propensity to file for bankruptcy would be to oversimplify the problem. Since the model accurately categorizes roughly 3 out of 4 to 9 out of 10 firms, thus not exclusively accurate classification, potential losses can be made by the creditor both in terms of granting credit to a unsound firm and fail to grant credit to a healthy firm. Although the results in this study are in line with prior research in terms of model prediction accuracy (Fletcher et al 1993; Yang et al 1999; Zhang et al 1999; Atiya 2001; Baek et al 2003) there are studies that produce prediction accuracies in a higher range of around 90 to 95% (Ohlson 1980; Back et al 1997).

CHAPTER 6
CONCLUSIONS

In the last chapter some short concluding remarks are presented along with possible improvements for further research within the field of bankruptcy prediction. A short remainder of the study is given followed by the relevancy of this study and some guidelines for the primary beneficiaries of this study.

This thesis examines the possibility to predict bankruptcy on Swedish corporations by using traditional quantitative methods (multiple discriminant analysis, logistic regression and neural networks). The obtained results indicate that all versions of logistic regression (standard, forward and backward conditional) that are applied in this study perform better than the other methods for the in-sample firms. The standard logistic regression also remains the best model for the holdout sample. However, the forward and backward conditional logistic regressions are outperformed by the neural network models (feedforward, backward and probabilistic). The highest prediction accuracy for the in-sample is obtained by backward conditional logistic regression (86,12%) and for the holdout sample it is the standard logistic regression (86,97%). We believe that neural network models would outperform the other methods if the data sample is larger. The two functional models (MDA and Logit) are easily applied while the neural network approach requires computational support. If the models are correctly applied a creditor can predict bankruptcy of up to 87%.

Traditional quantitative methods can be applied when a lender is assessing the bankruptcy probability of a borrower. However, we believe that the traditional quantitative methods should be applied as a complementary tool because they do not predict corporate bankruptcy with 100% certainty. Moreover, when a corporation applies for a loan then the borrowing amount is usually large and since there is a probability of a wrong prediction, a lender should not depend solely on traditional quantitative methods. Finally, this thesis provides the creditor with a valuable tool in assessing the propensity of bankruptcy for a corporate client. All models provided in this study can benefit a corporate creditor when assessing the probability of a corporate loan applicant. However, the best model for predicting corporate bankruptcy should be employed, which is the standard logistic regression in its functional form that is given in this study (see section “4.4.1 Standard Model”).

Primary beneficiaries of this study, such as large finance institutes and banks, should thus implement the following approach: (i) Consider the quantitative risk analysis as a complementary tool and; (ii) Apply the standard logistic model in its functional form (given in this study) to predict the probability of corporate bankruptcy for an individual corporate client when they receive a loan application.

Further research in this area could put a focus on gathering more data in order to make the neural networks perform better. Also, a different approach in pre-processing the data could yield another and more accurate result.

REFERENCES

LITERATURE

Afifi A; Clark V (1990) *"Computer-aided multivariate analysis"*. Second Edition, Van Nostrand Reinhold Company, New York, USA.

Kim K; Nofsinger J (2004) *"Corporate Governance"* International Edition, Pearson Prentice Hall, New Jersey, USA.

Kutner M; Nachtsheim C; Neter J (2004) *"Applied Linear Regression Models"*. Fourth Edition, McGraw-Hill Irwin, New York, USA.

ARTICLES

Altman E (1968) "The prediction of corporate bankruptcy: A discriminant analysis". *American Finance Association*, **23:1:** 193-194.

Altman E (1968) "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy". *American Finance Association*, **23:4:** 589-609.

Altman E (1984) "The success of business failure prediction models: An international survey". *Journal of Banking and Finance*, **8:2:** 171-199.

Altman E; Marco G; Varetto F (1994) "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)". *Journal of Banking and Finance*, **18:3:** 505-529.

Altman E (2000) "Predicting financial distress of companies: Revisiting the z-score and zeta models". *Citeseer*.

Anscombe F J (1967) "Topics in the investigation of linear relations fitted by the method of least squares". *Journal of the Royal Statistical Society*, **29:1:** 1-52.

Arshad K (1985) "Analyzing financial statements for managerial performance measurement and bankruptcy prediction". *Engineering Management International*, **3:3:** 165-175.

Atiya A (2001) "Bankruptcy prediction for credit risk using neural networks: A survey and new results". *IEEE Transactions on Neural Networks*, **12:4:** 929-935.

Aziz A M; Dar Humayon A (2006) "Predicting corporate bankruptcy: where we stand?". *Corporate Governance*, **6:1:** 18-33.

Back B; Laitinen T; Sere K; Van Wezel M (1997) “Choosing bankruptcy predictors using discriminant analysis, logit analysis, and genetic algorithms”. *Citeseer*.

Baek J; Cho S (2003) “Bankruptcy prediction for credit risk using an auto-associative neural network in korean firms”. *Computational Intelligence for Financial Engineering*, 25-29.

Barnes P (1982) “Methodological implications of non-normally distributed financial ratios”. *Journal of Business Finance & Accounting*, **9:1:** 51.

Barniv R; Agarwal A; Leach R (1997) “Predicting the outcome following bankruptcy filing: a three-state classification using neural networks”. *International Journal of Intelligent Systems in Accounting, Finance & Management*, **6:3:** 177-194.

Beaulieu P (1996) “A note on the role of memory in commercial loan officers’ use of accounting and character information”. *Accounting, Organizations and Society*, **21:6:** 515-528.

Beaver W (1966) “Financial ratios as predictors of failure”. *Journal of Accounting Research*, **4:** 71-111.

Becchetti L; Sierra J (2003) “Bankruptcy risk and productive efficiency in manufacturing firms”. *Journal of Banking and Finance*, **27:11:** 2099-2120.

Boritz E; Kennedy B (1995) “Effectiveness of neural network yypes for prediction of business failure”. *Expert Systems with Applications*, **9:4:** 503-512.

Charalambous C; Charitou A; Kaourou F (2000) “Comparative analysis of artificial neural network models: application in bankruptcy prediction”. *Annals of Operational Research*, **99:1-4:** 403-425.

Chi L C; Tang T C (2006) ”Bankruptcy prediction: application of logit analysis in export credit risks”. *Australian Journal of Management*, **31:1:** 17-27.

Cielen A; Peeters L & Vanhoof K (2004) “Bankruptcy prediction using a data envelopment analysis”. *European Journal of Operational Research*, **154:2:** 526-532.

Cox D R (1972) “The analysis of multivariate binary data”. *Applied Statistics*, **21:2:** 113-120.

Dietrich J; Arcelus F J; Srinivasan G (2005) “Predicting financial failure: Some evidence from New Brunswick agricultural co-ops”. *Annals of Public and Cooperative Economics*, **76:2:** 179-194.

Dugan M & Zavgren C (1989) “How a bankruptcy model could be incorporated as an analytical procedure”. *The CPA Journal*, **59:5:** 64-65.

Efron B (1975) “The efficiency of logistic regression compared to normal discriminant analysis”. *Journal of the American Statistical Association*, **70:352:** 892-898.

Eisenbeis R (1977) “Pitfalls in the application of discriminant analysis in business, finance, and economics”. *The Journal of Finance*, **32:3:** 875-900.

Etheridge H; Sriram R (1997) “A comparison of the relative costs of financial distress models: artificial neural networks, logit and multivariate discriminant analysis”. *International Journal of Intelligent Systems in Accounting, Finance & Management*, **6:3:** 235-248.

Fletcher D; Goss E (1993) “Forecasting with neural networks - An application using bankruptcy data”. *Information and Management*, **24:3:** 159-167.

Gaganis C; Pasiouras F; Doumpos M (2007) “Probabilistic neural networks for the identification of qualified audit opinions”. *Expert Systems with Applications*, **32:** 114-124.

Gong G (1986) “Cross-Validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression”. *Journal of the American Statistical Association*, **81:393:** 108-113.

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

Guttler A; Wahrenburg M (2007) “The adjustment of credit ratings in advance of defaults”. *Journal of Banking and Finance*, **31:3:** 751-767.

Harvey A C (1977) “Some comments on multicollinearity in regression”. *Applied Statistics*, **26:2:** 188-191.

Hillegeist S; Keating E; Cram D; Lundstedt K (2004) “Assessing the probability of bankruptcy”. *Review of Accounting Studies*, **9:1:** 5-34.

Holmen J (1988) “Using financial ratios to predict bankruptcy: An evaluation of classic models using recent evidence”. *Akron Business and Economic Review*, **19:1:** 52-64.

Hua Z; Wang Y; Xu X; Zhang B; Liang L (2007) “Predicting corporate financial distress based on integration of support vector machine and logistic regression”. *Expert Systems with Applications*, **33:2:** 434-440.

Johnsen T; Melicher R (1994) “Predicting corporate bankruptcy and financial distress: Information value added by multinomial logit models”. *Journal of Economics and Business*, **46:4:** 269-287.

Joy M; Tollefson J (1975) “On the financial applications of discriminant analysis”. *Journal of Financial and Quantitative Analysis*, **10:5:** 723-739.

Kim H; Gu Z (2006) “Predicting Restaurant Bankruptcy: A Logit Model in Comparison with a Discriminant Model”. *Journal of Hospitality and Tourism Research*, **30:4:** 474-493.

Kisgen D (2006) “Credit ratings and capital structure”. *The Journal of Finance*, **61:3:** 1035-1072.

Lachenbruch P (1968) “On expected probabilities of misclassification in discriminant analysis, necessary sample size, and a relation with the multiple correlation coefficient”. *Biometrics*, **24:4:** 823-834.

Lachenbruch P (1976) “Discriminant analysis”. *Journal of the Royal Statistical Society*, **139:2:** 270-271.

Laitinen E; Laitinen T (2000) “Bankruptcy prediction: Application of the Taylor's expansion in logistic regression”. *International Review of Financial Analysis*, **9:4:** 327-349.

Leshno M; Spector Y (1996) “Neural network prediction analysis: The bankruptcy case”. *Neurocomputing*, **10:2:** 125-147.

Lo A (1985) “Logit versus discriminant analysis: A specification test and application to corporate bankruptcies”. *Journal of Econometrics*, **31:1:** 151-178.

Mantel N; Brown C (1974) “Alternative test for comparing normal distribution parameters based on logistic regression”. *Biometrics*, **30:3:** 485-497.

Mensah Y (1984) “An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study”. *Journal of Accounting Research*, **22:1:** 380-395.

Nishikawa T (2002) “Credit risk estimation for small businesses based on a statistical method”. *Asia-Pacific Financial Markets*, **9:2:** 101-126.

Ohlson J (1980) “Financial ratios and the probabilistic prediction of bankruptcy”. *Journal of Accounting Research*, **18:1:** 109-131.

Pompe P; Bilderbeek J (2005) “The prediction of bankruptcy of small- and medium-sized industrial firms”. *Journal of Business Venturing*, **20:6:** 847-868.

Pompe P; Bilderbeek J (2005) “Bankruptcy prediction: the influence of the year prior to failure selected for model building and the effects in a period of economic decline”. *International Journal of Intelligent Systems in Accounting, Finance and Management*, **13:2:** 95-112.

Raghupathi W; Schkade L; Raju B (1991) “A neural network application for bankruptcy prediction”. *System Sciences, 1991, Proceedings of the Twenty-Fourth Annual Hawaii International Conference*, **4:** 147-155.

Rosenberg E; Gleit A (1994) “Quantitative methods in credit management: A survey”. *Operations research*, **42:4:** 589-613.

Salam F; Bai S (1993) “A feedback neural network with supervised learning”. Neural Networks, IEEE Transactions, **2:1:** 170-173.

Scott J (1981) “The probability of bankruptcy: A comparison of empirical predictions and theoretical models”. *Journal of Banking and Finance*, **5:3:** 317-345.

Serrano-Cinca C (1997) “Feedforward neural networks in the classification of financial information”. *The European Journal of Finance*, **3:3:** 183-202.

Sheppard J (1994) “The dilemma of matched pairs and diversified firms in bankruptcy prediction models”. *The Mid-Atlantic Journal of Business*, **30:1:** 9-25.

Spanos M; Dounias G; Matsatsinis N; Zopounidis C (2001) “A fuzzy knowledge-based decision aiding method for the assessment of financial risks; The case of corporate bankruptcy prediction”. *CiteSeer*.

Sudheer C; Jarrow R (2004) “Bankruptcy prediction with industry effects”. *Review of Finance*, **8:4:** 537-567.

Tegerstedt G; Miedel M; Maehle-Schmidt M; Nyren O; Hammarström (2005) ”A short-form questionnaire identified genital organ prolapse”. *Journal of Clinical Epidemology*, **58:1:** 41-46.

Tucker J (1996) “Neural networks versus logistic regression in financial modelling: A methodological comparison”. *Citeseer*.

Warner B; Manavendra M (1970) “Understanding neural networks as statistical tools”. *Citeseer*.

Welch B L (1939) “Note on discriminant analysis”. *Biometrika*, **31:1/2:** 218-220.

White H (1980) “A heteroskedasticity-vonsistent covariance matrix estimator and a direct test for heteroskedasticity”. *Econometrica: The Journal of the Econometric Society*, **48:4:** 817-838.

Wilson R; Sharda R (1994) “Bankruptcy prediction using neural networks”. *Decision Support Systems*, **11:5:** 545-557.

Yang Z; Platt M; Platt H (1999) “Probabilistic neural networks in bankruptcy prediction”. *Journal of Business Research*, 44:2: 67-74.

Young R; Yue W T (2005) “Firm bankruptcy prediction; Experimental comparison of isotonic separation and other classification approaches”. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, **35:5:** 727-737.

Zavgren C (1985) “Assessing the vulnerability to failure of American industrial firms: A logistic analysis”. *Journal of Business Finance and Accounting*, **12:1:** 1-19.

Zhang G; Hu M; Patuwo E; Indro D (1999) “Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis”. *European Journal of Operational Research*, **116:1:** 16-32.

INTERNET

http://www.riskglossary.com/link/credit_risk.htm

http://www.riskglossary.com/link/credit_risk.htm

http://www.sbrn.org/Connections/10_03_Commercial_Lending.htm

http://www.bizjournals.com/birmingham/stories/2002/07/01/focus6.html?page=1

http://www.ansci.cornell.edu/tmplobs/doc77.pdf

http://www.ansci.cornell.edu/tmplobs/doc77.pdf

http://www.loanuniverse.com/credit.html

http://www.ansci.cornell.edu/tmplobs/doc77.pdf

http://www.ansci.cornell.edu/tmplobs/doc77.pdf

http://www.loanuniverse.com/credit.html

http://www.loanuniverse.com/credit.html

http://www.ansci.cornell.edu/tmplobs/doc77.pdf

http://149.170.199.144/multivar/dawords.htm

http://faculty.vassar.edu/lowry/lr1.gif

DATABASES

Affärsdata	www.ad.se
ELIN (Electronic Library Information Navigator)	http://elin.lub.lu.se.ludwig.lub.lu.se/elin
SCB (Statistiska Centralbyrån)	www.scb.se
LIBRIS (National database for library services)	

APPENDIX 1

AGE OF FIRM BY NUMBERS AND PERCENTAGES

Age Interval (years)	*Defaulted Firms*	*Non-defaulted Firms*
0 – 5	21	8
6 – 10	28	22
11 – 20	39	38
21 – 30	9	13
31 – 50	12	21
51 – 100	4	10
100 –	0	1
TOTAL	**113**	**113**
0 – 5	18,6%	7,1%
6 – 10	24,8%	19,5%
11 – 20	34,5%	33,6%
21 – 30	8,0%	11,5%
31 – 50	10,6%	18,6%
51 – 100	3,5%	8,8%
100 –	0%	0,9%
TOTAL	**100,0%**	**100,0%**

APPENDIX 1: TABLE 1. Firm according to age.

APPENDIX 2

SIZE OF FIRM MEASURED BY EMPLOYEES IN NUMBERS AND PERCENTAGES

Number of employees	*Defaulted Firms*	*Non-defaulted Firms*
30 – 35	34	22
36 – 40	15	18
41 – 50	26	19
51 – 75	22	26
76 – 100	9	8
101 – 150	5	9
151 – 250	1	8
250 –	1	3
TOTAL	**113**	**113**
30 – 35	30,0%	19,4%
36 – 40	13,3%	16,0%
41 – 50	23,0%	16,8%
51 – 75	19,5%	23,0%
76 – 100	8,0%	7,1%
101 – 150	4,4%	8,0%
151 – 250	0,9%	7,1%
250 –	0,9%	2,6%
TOTAL	**100,0%**	**100,0%**

APPENDIX 2: TABLE 1. Firm size according to the number of employees.

APPENDIX 3

Matches

IN-SAMPLE FIRMS

Defaulted firm (D1 to D90)	Matched firm (N1 to N90)	Year
Apport PostFoto AB	Optoteam AB	2004
Benefit Karlstad AB	Conagri AB	2003
ETC Produktion AB	Ingress Media AB	2005
Electroprocess i Dalarna AB	Instrumentfirman Inor Process AB	2004
Ardeo Telemarketing AB	Online i Lund AB	2005
ATAP Protection Group AB	Rapid Bevakning AB	2006
Alfta Rehab Center AB	Medic Operating Support AB	2006
Alpha Sweden AB	Fehrer Sweden AB	2006
Anhersson AB	VÖFAB	2005
Animonhus AB	A-Hus AB	2005
BELAB	Rivners AB	2005
BK Tåg AB	TGOJ Trafik AB	2005
BSP Svets & Plåt AB	Fromells Vip-Teknik AB	2004
Bankeryd Maskin AB	Ventur Tekniska AB	2005
Bilgallerian Thomas Jönsson AB	Ambjörnssons Bil i Jönköping AB	2005
Bjurfors Bygg AB	Aros Bygg & Förvaltning AB	2005
Boat Sales International Marine AB	Swede Ship Marine AB	2005
Call & Research no1 AB	Hermelin Nordic Research AB	2005
Call4u AB	Teleresurs i Sverige AB	2005
Care Service BM AB	Mercordia AB	2006
Centrifug Data & Handel AB	Autodesk AB	2005
Charkmästarn Västerås AB	Direkt Chark i Göteborg AB	2006
Cidema AB	Mecel AB	2004
Componenta Alvesta AB	Roslagsgjuteriet AB	2006
Connection Consulting Sverige AB	Infrasystems Solutions Stockholm AB	2005
Crew Bemanning Västsverige AB	AMENDO Bemanning AB	2005
D Cultmag AB	Comcarta AB	2005
DB Grafiska AB	Holmbergs i Malmö AB	2005
Distributören i Gävle AB	Randstad AB	2006
ETK Mur & Puts AB	Puts & Tegel i Örebro AB	2007
EkSec AB	Gestrike Vakt AB	2005
Elektrotryck AB	Råda Elektriska AB	2006
Envirocleaners Sverige AB	Mälarsanering Sverige AB	2006
Exicometer Consulting AB	LinPro AB	2007
FA-Tec AB	Hörle Automatic AB	2006
FCC Sprinkler AB	Olsbergs Hydraulics AB	2004
Falcon Air AB	AB Norrlandsflyg	2006
Figal AB	AB Maskinarbeten	2006
Forest King Sawmill AB	Svenska Skogsplantor AB	2005
Form3 interior AB	EM Home Interior AB	2006
Foto Design i Mullsjö AB	Big Image Systems, Sweden AB	2007
G-Lack Automotive AB	Forslunds Bil AB	2006
G.U. Consulting AB	Iris Utvecklingscenter AB	2005
Gavle Marknadskommunikation AB	Scandinavian Human Resources AB	2004

Gjuteribolaget i Bredaryd AB	Nomeko AB	2006
Gudö Puts och Montage AB	Harrysson Entreprenad AB (HEAB)	2005
Göta Timmele Färgeri AB	Sjuhäradsbygdens Färgeri AB	2006
Göteborgs AB Scanauto	Roy Andersson Bilbolaget AB	2005
Hedemora Verkstäder AB	Mjöbäcks Entreprenad AB	2004
Hultafors Hälsocenter AB	EF Medical AB	2005
Hydraul & Mekan i Lappland AB	Laholm Industrier AB	2006
I.B. Mekaniska i Kungälv AB	MARINFLOC AB	2006
In or Out International AB	Iris Hantverk AB	2006
Intensivutbildning Sverige AB	UVS Utbildning AB	2006
Johns Bygg & Fasad, John Fellert AB	Berg & Byggteknik i Norberg AB	2004
JMS Tidningstryckarna AB	Borås Tidning Tryckeri AB	2006
Klippan AB	J D Stenqvist AB	2006
K2 Trainpool AB	ARF Netcom AB	2005
Kallanders Åkeri AB	Landet Runt i Strängnäs AB	2005
Klippan Mölndal AB	Eson-Pac AB	2006
Kvalitetsproduktion i Degerfors AB	Brandon AB	2006
Lappmarkshotell AB	Hotell Statt i Eskilstuna AB	2005
Laser Tool i Valdemarsvik AB	Liljeholmens Stearinfabriks AB	2006
Licenssvets Rör i Sölvesborg AB	ES Traffic AB	2005
Lindman Telecom AB	Teknova Byggsystem AB	2004
Ljud@Bild Digital Center Svenska AB	Din ELON butik AB	2005
Ljungföretagen Tryckeri AB	Stora Enso Hylte AB	2004
Lokalvårdarna Hernandez & Co AB	Rengörare Näslund AB	2006
MKPK HANTVERKSSERVICE AB	LC Service & Renova i Sverige AB	2005
M. Wennerstens Åkeri AB	Hamburg Süd Norden AB	2006
M.A.S.A.B. Magnus Strömberg AB	Forsgrens Timmerhus AB	2004
Maximiliam AB	Kumlins Måleri i Umeå AB	2006
Nordic East Airlink AB	Stockholm Fuelling Services AB	2004
Nordisk Industri- och VVS-service AB	Arena Personal Stockholm AB	2005
Northlight Optronics AB	Glocalnet AB	2006
OctaCom Stockholm AB	Spray Network AB	2006
Personalhantering i Lunda AB	Personalhantering i Lunda AB	2005
Proteus Proffs assistans AB	Firesafe Sverige AB	2006
Ramsåsa Åkeri AB	Skandinavisk Gränsspedition AB	2005
Remarkable Retail Marketing AB	AGRI Marketing AB	2006
Restaurang BDM AB	Helnan International Hotels AB	2006
Rotik AB	European Nursery Group Sweden AB	2006
Rulles Fruktlåda i Tumba AB	Fruktbudet i Stockholm AB	2005
SMA Construction AB	Tage Rejmes i Örebro Lastvagnar AB	2004
SMA Maskin AB	Lecab Bilaktiebolag	2004
SMT Tricept AB	Know IT Candeo AB	2004
SPR Svenska Personal AB	Expandera Mera AB	2006
Sandberg & Trygg AB	Hemfrid i Sverige AB	2006
SE Bygg & Förvaltning AB	SSC Skellefteå AB	2006
Skandinaviska Sprinkler AB	ES Traffic AB	2005

HOLDOUT-SAMPLE FIRMS

Defaulted firm (d1 to d23)	**Matched firm (n1 to n23)**	**Year**
Skivhugget AB	Akademibokhandelgruppen AB	2004
Solstatrafik Fjärr AB	June Express AB	2005
Staben i Tungelsta AB	HomeMaid Hemservice AB	2007
Stenberg & Co Måleri AB	Engbergs Måleri i Uppsala AB	2006

Svensk Kompositutveckling AB	Götene Plast AB	2006
Svetsbröderna i Sveg, Örebro AB	Göteborgs Smörjmedelsfabrik AB	2004
Swedline Express AB	AFCO AB	2006
Swedtel AB	COLT Telecom AB	2007
Swesco Bolander AB	AB Svenskt Konstsilke	2004
Sydkött AB	Team Ugglarp AB	2005
T P Promotion AB	ISS Demogruppen AB	2005
TRISAB i Norr AB	Bröderna Magnusson Transport AB	2005
TRISAB i Syd AB	Jetpak Sverige AB	2005
Techsite International AB	Axiell Bibliotek AB	2004
Tidningsdistribution i Göteborg AB	Hemfrid i Sverige AB	2005
Tivox Automation AB	Strömma Fjäll & Aktivitet AB	2005
Transfixx Örebro AB	CombiTrans Groupage AB	2006
Vacnox AB	Nordic Alarm AB	2006
VBS i Oskarshamn AB	Iris Hantverk AB	2004
Västrumsgården AB	Capio Artro Clinic AB	2005
Workman i Stockholm AB	Actit AB	2004
Ytel AB	Access Personnel AB	2005
by Kjellberg AB	Harrysson Entreprenad AB (HEAB)	2006

APPENDIX 4 (See further APPENDIX 3 for specification of firm)

D = Defaulted in-sample firm

N = Non-defaulted in-sample firm

d = Defaulted holdout-sample firm

n = Non-defaulted holdout-sample firm

IN-SAMPLE FIRMS	SCORES						
	MDA	*Standard Logit*	*Forward Logit*	*Backward Logit*	*Feedforward NN*	*Feedback NN*	*Prob. NN*
D1	1,512	1,000	1,000	1,000	1,000	1,000	1,000
D2	0,513	0,989	0,471	0,959	0,542	0,542	0,542
D3	0,593	0,936	0,487	0,730	0,403	0,403	0,403
D4	0,693	0,899	0,831	0,911	0,735	0,735	0,735
D5	1,142	0,966	0,932	0,985	0,832	0,832	0,832
D6	2,200	0,944	0,974	0,987	0,897	0,897	0,897
D7	1,145	1,000	0,995	1,000	0,947	0,947	0,947
D8	-0,156	0,102	0,868	0,729	0,787	0,787	0,787
D9	0,743	0,629	0,782	0,787	0,610	0,610	0,610
D10	0,204	0,477	0,791	0,757	0,551	0,551	0,551
D11	0,940	0,480	0,470	0,678	0,445	0,445	0,445
D12	-0,178	0,396	0,300	0,413	0,229	0,229	0,229
D13	0,663	0,591	0,770	0,879	0,579	0,579	0,579
D14	0,646	0,698	0,908	0,949	0,721	0,721	0,721
D15	-0,088	0,396	0,559	0,574	0,342	0,342	0,342
D16	0,615	0,589	0,708	0,639	0,642	0,642	0,642
D17	1,407	1,000	1,000	1,000	1,000	1,000	1,000
D18	1,944	1,000	1,000	1,000	1,000	1,000	1,000
D19	1,338	0,999	0,999	1,000	0,999	0,999	0,999
D20	0,355	0,141	0,187	0,099	0,341	0,341	0,341
D21	0,669	1,000	0,708	1,000	1,000	1,000	1,000
D22	0,463	0,978	0,987	0,996	0,858	0,858	0,858
D23	0,593	0,547	0,394	0,364	0,517	0,517	0,517
D24	0,996	1,000	0,999	1,000	0,992	0,992	0,992
D25	1,067	0,982	0,793	0,957	0,734	0,734	0,734
D26	0,840	0,915	0,979	0,989	0,748	0,748	0,748
D27	1,001	0,963	0,528	0,739	0,613	0,613	0,613
D28	0,606	0,756	0,831	0,901	0,656	0,656	0,656
D29	0,242	0,638	0,345	0,449	0,517	0,517	0,517
D30	0,562	0,651	0,217	0,482	0,266	0,266	0,266
D31	0,555	0,993	0,944	1,000	0,863	0,863	0,863
D32	0,437	0,607	0,656	0,631	0,472	0,472	0,472
D33	0,955	0,974	0,929	0,975	0,859	0,859	0,859
D34	-0,309	0,055	0,110	0,188	0,158	0,158	0,158
D35	0,763	0,684	0,916	0,937	0,770	0,770	0,770
D36	0,441	0,877	0,774	0,916	0,575	0,575	0,575
D37	-0,478	0,271	0,089	0,081	0,042	0,042	0,042
D38	1,255	1,000	1,000	1,000	1,000	1,000	1,000

D39	0,438	0,993	0,780	0,745	0,781	0,781	0,781
D40	0,867	0,605	0,804	0,795	0,635	0,635	0,635
D41	0,675	0,405	0,307	0,533	0,323	0,323	0,323
D42	0,771	0,693	0,749	0,807	0,731	0,731	0,731
D43	2,187	1,000	1,000	1,000	1,000	1,000	1,000
D44	1,419	0,998	0,999	1,000	1,000	1,000	1,000
D45	0,793	0,907	0,914	0,897	0,653	0,653	0,653
D46	0,748	0,783	0,811	0,786	0,620	0,620	0,620
D47	0,754	0,793	0,798	0,963	0,676	0,676	0,676
D48	0,206	0,892	0,960	0,962	0,636	0,636	0,636
D49	0,403	0,615	0,645	0,497	0,668	0,668	0,668
D50	0,903	0,916	0,809	0,976	0,512	0,512	0,512
D51	0,820	0,869	0,978	0,988	0,731	0,731	0,731
D52	0,633	0,832	0,772	0,821	0,661	0,661	0,661
D53	0,698	0,719	0,733	0,768	0,545	0,545	0,545
D54	0,833	0,738	0,594	0,468	0,786	0,786	0,786
D55	0,587	0,410	0,634	0,401	0,544	0,544	0,544
D56	0,045	0,413	0,292	0,541	0,238	0,238	0,238
D57	-0,842	0,387	0,192	0,293	0,214	0,214	0,214
D58	1,474	0,902	0,929	0,950	0,821	0,821	0,821
D59	0,843	0,579	0,850	0,722	0,617	0,617	0,617
D60	-0,606	0,482	0,361	0,234	0,099	0,099	0,099
D61	0,401	0,504	0,802	0,924	0,624	0,624	0,624
D62	0,878	0,782	0,752	0,838	0,774	0,774	0,774
D63	0,827	0,762	0,858	0,768	0,489	0,489	0,489
D64	0,313	0,540	0,667	0,676	0,469	0,469	0,469
D65	0,115	0,332	0,467	0,653	0,392	0,392	0,392
D66	0,670	0,775	0,990	0,994	0,857	0,857	0,857
D67	0,572	0,854	0,945	0,966	0,762	0,762	0,762
D68	0,503	0,697	0,722	0,756	0,400	0,400	0,400
D69	0,582	0,788	0,613	0,765	0,486	0,486	0,486
D70	0,920	0,950	0,971	0,991	0,870	0,870	0,870
D71	0,825	0,932	0,779	0,809	0,561	0,561	0,561
D72	1,667	0,288	0,766	0,601	0,590	0,590	0,590
D73	1,621	1,000	1,000	1,000	1,000	1,000	1,000
D74	0,550	0,555	0,588	0,491	0,392	0,392	0,392
D75	1,262	1,000	0,887	0,993	0,933	0,933	0,933
D76	0,693	0,729	0,682	0,580	0,666	0,666	0,666
D77	1,386	1,000	1,000	1,000	1,000	1,000	1,000
D78	1,273	0,978	0,994	0,980	0,986	0,986	0,986
D79	0,667	0,978	0,990	0,999	0,829	0,829	0,829
D80	0,702	0,778	0,751	0,892	0,527	0,527	0,527
D81	3,077	1,000	1,000	0,959	1,000	1,000	1,000
D82	0,482	0,563	0,783	0,715	0,551	0,551	0,551
D83	-0,378	0,577	0,439	0,592	0,310	0,310	0,310
D84	-0,166	0,698	0,761	0,740	0,484	0,484	0,484
D85	0,065	0,696	0,818	0,773	0,490	0,490	0,490
D86	0,538	1,000	0,959	0,998	0,917	0,917	0,917
D87	1,216	0,877	0,846	0,875	0,604	0,604	0,604
D88	0,309	0,446	0,295	0,789	0,327	0,327	0,327
D89	-0,011	0,394	0,435	0,108	0,165	0,165	0,165
D90	0,632	0,589	0,798	0,542	0,416	0,416	0,416
N1	0,085	0,330	0,409	0,378	0,394	0,394	0,394
N2	-0,117	0,397	0,420	0,367	0,283	0,283	0,283

N3	-0,140	0,093	0,168	0,008	0,000	0,000	0,000
N4	0,318	0,345	0,546	0,645	0,494	0,494	0,494
N5	-0,368	0,018	0,010	0,013	0,000	0,000	0,000
N6	0,453	0,498	0,650	0,665	0,487	0,487	0,487
N7	-0,230	0,227	0,331	0,383	0,189	0,189	0,189
N8	-0,570	0,108	0,095	0,046	0,067	0,067	0,067
N9	-0,104	0,637	0,260	0,396	0,322	0,322	0,322
N10	-0,493	0,085	0,203	0,136	0,172	0,172	0,172
N11	-0,710	0,012	0,003	0,006	0,000	0,000	0,000
N12	-0,690	0,131	0,090	0,069	0,097	0,097	0,097
N13	0,602	0,594	0,734	0,776	0,571	0,571	0,571
N14	-0,393	0,012	0,010	0,006	0,000	0,000	0,000
N15	0,216	0,590	0,658	0,649	0,543	0,543	0,543
N16	0,139	0,415	0,442	0,448	0,372	0,372	0,372
N17	-0,693	0,154	0,045	0,059	0,001	0,001	0,001
N18	0,040	0,241	0,237	0,156	0,296	0,296	0,296
N19	0,427	0,143	0,171	0,071	0,147	0,147	0,147
N20	-0,189	0,037	0,095	0,061	0,109	0,109	0,109
N21	-0,071	0,022	0,017	0,022	0,056	0,056	0,056
N22	-0,405	0,068	0,036	0,013	0,071	0,071	0,071
N23	-0,402	0,067	0,043	0,063	0,009	0,009	0,009
N24	-0,653	0,182	0,094	0,031	0,031	0,031	0,031
N25	-0,146	0,113	0,126	0,034	0,240	0,240	0,240
N26	0,291	0,150	0,134	0,058	0,162	0,162	0,162
N27	-0,178	0,101	0,222	0,232	0,187	0,187	0,187
N28	-0,175	0,119	0,056	0,008	0,000	0,000	0,000
N29	-0,205	0,031	0,056	0,028	0,019	0,019	0,019
N30	-0,154	0,120	0,085	0,099	0,175	0,175	0,175
N31	0,800	0,512	0,616	0,485	0,417	0,417	0,417
N32	-0,210	0,015	0,009	0,019	0,000	0,000	0,000
N33	-0,432	0,039	0,010	0,007	0,000	0,000	0,000
N34	0,455	0,202	0,355	0,170	0,302	0,302	0,302
N35	-0,356	0,175	0,107	0,135	0,132	0,132	0,132
N36	-1,038	0,031	0,004	0,015	0,000	0,000	0,000
N37	0,212	0,520	0,564	0,387	0,462	0,462	0,462
N38	-0,758	0,006	0,005	0,011	0,000	0,000	0,000
N39	-0,336	0,100	0,484	0,609	0,461	0,461	0,461
N40	-0,761	0,331	0,219	0,220	0,000	0,000	0,000
N41	-0,059	0,057	0,026	0,021	0,000	0,000	0,000
N42	-0,719	0,211	0,116	0,178	0,194	0,194	0,194
N43	0,549	0,445	0,555	0,631	0,546	0,546	0,546
N44	0,641	0,580	0,057	0,107	0,083	0,083	0,083
N45	-0,602	0,117	0,064	0,027	0,048	0,048	0,048
N46	-0,349	0,132	0,032	0,076	0,079	0,079	0,079
N47	0,243	0,519	0,478	0,481	0,455	0,455	0,455
N48	-0,692	0,144	0,104	0,056	0,110	0,110	0,110
N49	-0,371	0,263	0,333	0,273	0,176	0,176	0,176
N50	-0,385	0,048	0,011	0,016	0,023	0,023	0,023
N51	0,466	0,183	0,230	0,231	0,266	0,266	0,266
N52	0,467	0,485	0,748	0,726	0,585	0,585	0,585
N53	0,408	0,338	0,269	0,347	0,222	0,222	0,222
N54	1,240	0,973	0,717	0,271	0,779	0,779	0,779
N55	0,123	0,266	0,179	0,332	0,290	0,290	0,290
N56	-0,055	0,436	0,484	0,458	0,281	0,281	0,281

N57	-1,302	0,033	0,006	0,002	0,000	0,000	0,000
N58	0,429	0,147	0,274	0,084	0,240	0,240	0,240
N59	0,635	0,696	0,413	0,379	0,541	0,541	0,541
N60	-0,772	0,116	0,034	0,032	0,038	0,038	0,038
N61	0,184	0,348	0,452	0,463	0,319	0,319	0,319
N62	0,272	0,557	0,239	0,154	0,461	0,461	0,461
N63	-0,743	0,149	0,005	0,004	0,000	0,000	0,000
N64	0,029	0,196	0,123	0,237	0,180	0,180	0,180
N65	0,139	0,348	0,556	0,451	0,418	0,418	0,418
N66	-0,006	0,271	0,511	0,386	0,404	0,404	0,404
N67	-2,570	0,063	0,004	0,005	0,000	0,000	0,000
N68	0,276	0,104	0,169	0,058	0,067	0,067	0,067
N69	1,334	0,987	0,997	0,997	1,000	1,000	1,000
N70	-0,035	0,365	0,063	0,019	0,000	0,000	0,000
N71	-0,309	0,092	0,040	0,049	0,124	0,124	0,124
N72	0,375	0,182	0,324	0,115	0,328	0,328	0,328
N73	0,740	0,732	0,474	0,500	0,521	0,521	0,521
N74	0,862	0,743	0,768	0,659	0,762	0,762	0,762
N75	-0,999	0,254	0,037	0,023	0,000	0,000	0,000
N76	-0,599	0,040	0,103	0,421	0,231	0,231	0,231
N77	0,389	0,135	0,075	0,039	0,285	0,285	0,285
N78	0,905	0,923	0,926	0,976	0,649	0,649	0,649
N79	0,609	0,111	0,484	0,093	0,286	0,286	0,286
N80	0,937	0,561	0,616	0,298	0,653	0,653	0,653
N81	0,535	0,872	0,559	0,488	0,219	0,219	0,219
N82	-0,414	0,056	0,123	0,095	0,076	0,076	0,076
N83	0,082	0,010	0,014	0,005	0,011	0,011	0,011
N84	-0,292	0,188	0,271	0,027	0,170	0,170	0,170
N85	-0,625	0,165	0,235	0,101	0,133	0,133	0,133
N86	-0,303	0,002	0,022	0,000	0,023	0,023	0,023
N87	0,845	0,309	0,603	0,482	0,297	0,297	0,297
N88	0,280	0,326	0,277	0,269	0,180	0,180	0,180
N89	-0,470	0,439	0,317	0,248	0,177	0,177	0,177
N90	0,029	0,196	0,123	0,237	0,180	0,180	0,180
HOLDOUT_FIRMS							
d1	1,355	1,000	1,000	1,000	1,000	1,000	1,000
d2	0,447	0,669	0,600	0,787	0,657	0,657	0,657
d3	0,529	0,052	0,149	0,014	0,288	0,288	0,288
d4	0,888	0,966	0,994	0,998	0,816	0,816	0,816
d5	0,701	0,857	0,737	0,844	0,637	0,637	0,637
d6	0,618	0,606	0,759	0,723	0,622	0,622	0,622
d7	1,511	1,000	1,000	1,000	1,000	1,000	1,000
d8	0,040	0,175	0,164	0,073	0,112	0,112	0,112
d9	0,655	0,581	0,913	0,765	0,790	0,790	0,790
d10	-0,181	0,718	0,328	0,483	0,498	0,498	0,498
d11	1,109	0,949	0,855	0,828	0,707	0,707	0,707
d12	-0,127	0,438	0,206	0,255	0,392	0,392	0,392
d13	0,292	0,717	0,193	0,180	0,510	0,510	0,510
d14	1,260	1,000	1,000	1,000	1,000	1,000	1,000
d15	0,796	0,862	0,710	0,831	0,814	0,814	0,814
d16	0,306	0,511	0,066	0,018	0,845	0,845	0,845
d17	1,727	0,986	0,993	0,998	0,928	0,928	0,928
d18	0,760	0,520	0,808	0,765	0,568	0,568	0,568

d19	1,029	0,942	0,914	0,937	0,706	0,706	0,706
d20	0,564	0,541	0,380	0,507	0,479	0,479	0,479
d21	0,963	0,974	0,857	0,887	0,886	0,886	0,886
d22	1,127	0,910	0,900	0,947	0,768	0,768	0,768
d23	1,207	0,999	0,999	1,000	1,000	1,000	1,000
n1	0,244	0,163	0,189	0,346	0,220	0,220	0,220
n2	0,049	0,326	0,198	0,200	0,218	0,218	0,218
n3	0,492	0,486	0,545	0,603	0,479	0,479	0,479
n4	0,090	0,075	0,055	0,054	0,038	0,038	0,038
n5	-0,155	0,303	0,169	0,131	0,166	0,166	0,166
n6	-0,489	0,330	0,432	0,420	0,304	0,304	0,304
n7	0,776	0,752	0,692	0,673	0,423	0,423	0,423
n8	-0,035	0,088	0,000	0,000	0,272	0,272	0,272
n9	-0,437	0,294	0,070	0,015	0,004	0,004	0,004
n10	0,021	0,432	0,530	0,511	0,508	0,508	0,508
n11	0,767	0,488	0,609	0,279	0,559	0,559	0,559
n12	-0,178	0,455	0,321	0,425	0,422	0,422	0,422
n13	-0,793	0,044	0,012	0,022	0,002	0,002	0,002
n14	-0,542	0,096	0,041	0,040	0,058	0,058	0,058
n15	0,365	0,473	0,382	0,399	0,235	0,235	0,235
n16	0,220	0,892	0,824	0,960	0,596	0,596	0,596
n17	-0,742	0,176	0,112	0,052	0,075	0,075	0,075
n18	0,209	0,491	0,371	0,406	0,410	0,410	0,410
n19	0,348	0,139	0,127	0,259	0,231	0,231	0,231
n20	0,101	0,138	0,011	0,001	0,319	0,319	0,319
n21	0,407	0,317	0,409	0,290	0,290	0,290	0,290
n22	1,283	0,998	0,988	1,000	0,829	0,829	0,829
n23	-0,382	0,127	0,020	0,057	0,059	0,059	0,059