ASSESSMENT OF CREDIT RISK IN THE NORWEGIAN BUSINESS SECTOR

Espen Sjøvoll[‡]

Norges Bank

August 12, 1999

Abstract

In this thesis, I present a model that measures credit risk in the Norwegian business sector, using firm bankruptcy as proxy for credit risk. Probit analysis, a discrete response model, is applied to micro level financial information from more than 500 000 observations from the period 1989-1998. Bankruptcies in the period 1995-1998 are used to develop the model, and bankruptcies in the period 1991-1993 are used for out of sample testing. A set of time-consistent indicators of bankruptcy is found by combining ideas from both the theory of industrial organisation and financial statement analysis. The results support the idea of a learning effect in companies. This effect is recognised with reduced risk of bankruptcy when observations are subject to age. Furthermore, the results indicate that debt and interest burden increase risk of bankruptcy, while equity decrease risk of bankruptcy. Real-estate companies generally have a lower risk, while restaurants generally have a higher risk.

Keywords: Bankruptcy, probit estimation, credit risk

JEL Classification: C13, G33

[‡] Norges Bank, The Central Bank of Norway. E-mail: espen.sjovoll@norges-bank.no

Preface

This thesis represents the end of my Cand. Polit. degree in economics at the University of Bergen and is the finishing line for more than six years of university studies. The topic is prediction of bankruptcies using accounting data. My first contact with the field of bankruptcy prediction was during my studies in Kiel, Germany, where Christopher Blevins and I wrote a term paper on bankruptcies in Germany. Returning from Germany, I started working on this thesis in the beginning of July 1998.

The Research department of Norges Bank supported the work on this thesis financially. I appreciated the opportunity I had to write my thesis in such an inspiring environment. I especially thank Bent Vale, Trond Eklund, Bjarne Gulbrandsen, Tor Reistabakk and Tore Anders Husebø, in Norges Bank, for all the help and advice they have given me. I also thank my advisors at the University of Bergen, Espen Bratberg and Kjell Vaage. Without their helpful comments and constructive criticisms, this thesis would not have contained the insights it does.

Finally, I wish to thank both of my parents for the support and patience they have had with me. Finding my way through the academic labyrinth was stressful at best, but in the end, I found my way. In addition, I am grateful to my girlfriend for helping me adjust to life in Oslo.

Oslo,12 august 1999

Espen Sjøvoll

Table of Contents:

"Bankruptcy is essentially a state-supervised system for breaking contracts that are mutually inconsistent and therefore, unenforceable."

Michael C. Jensen, Presidential address to the American Finance Association 1993

1 INTRODUCTION

The goal of this thesis is to develop a model that measures credit risk in the Norwegian business sector by investigating corporate bankruptcy. In this thesis, consistent indicators of bankruptcy are found. This is done by combining ideas from both the theory of industrial organisation and financial statement analysis. The econometric method of the thesis is probit analysis, this method is applied on micro level financial information from more than 500 000 observations[1] from the period 1989-1998.

The motivation is to use micro economic information to determine the risk for bankruptcy. Changes in credit risk over time may provide important information on the development of the business cycle. Credit risk models are used to varying degrees in different countries and institutions. The Basle Committee on Banking Supervision[2] is currently surveying the extent of credit risk models in the banking sector. The results of this survey were not available at the time of writing[3]. Norges Bank (The Norwegian Central Bank) currently employs a model for risk analysis of the business sector as a part of its tasks of monitoring financial-system stability in Norway. In the thesis, an alternative model for measuring credit risk in the Norwegian business sector is proposed. Two academic fields treat this type of problems, and this thesis is an attempt to merge ideas from both of these fields.

The first field is theory of industrial organisation. Company exit is one of the phenomena that the theory of industrial organisation investigates. While exit is a wider term than bankruptcy, it is expected that some of the driving factors for company exit are applicable for the narrower group that bankruptcies are. Theory of industrial organisations looks at the company and its relation to the marketplace. Much of the focus is on company entry rather than company exit. This thesis incorporates ideas from: The vintage capital of Johansen (1959), the selection model of Jovanovic (1982), the exit model of Ghemewat and Nalebuff (1985) and theories of

[1] 507 880 observations are used, divided into an estimation sample of 322 842 observations, and a test sample of 185 038 observations.

[2] A part of the Bank for International Settlements

[3] The results of the study are expected to come out during the first half of 1999, see BIS (1998) for more information on the project.

recessionary cleansing of productivity proposed by Aghion and Howitt (1992) and Caballero and Hammour (1994, 1996).

The other field is financial statement analysis. Assessment of the risk of default is one of the focuses of financial statement analysis. The goal is normally to maximise the number of good loans and minimise the number of bad loans. The result is credit scoring models that can guide a creditor in the credit application decision process. The area of financial statement analysis I focus on is studies of financial distress among companies. Previous studies that compare the performance of different companies have found that the annual report of a company contains information that can indicate the risk of financial distress. The foundations for this type of model are, among others: Beaver (1966), Altman (1968), Altman *et al.* (1977), Wilcox (1971,1976), and Ohlson (1980). The impression of this work is that there is a lack of generality and agreement in terms of findings and methods.

To my knowledge, there have been no specific attempts to merge the ideas of the two different fields. On the one hand, there is the theory of industrial organisations that explains the possible mechanisms in a company and the differences between survivors and non-survivors. On the other hand, financial statement analysis attempts to quantify observable mechanisms. In this thesis, it is found support for the idea that these two fields complement each other.

Bankruptcies became an increasing problem in Norway in the early 1990s. Abolishment of credit market restrictions in the mid-1980s led to a high growth of company indebtedness. The recession in Norway that started in 1989/90[4] led to massive credit losses in Norwegian banks[5]. Table 1 shows changes in the level of bankruptcies in Norway, with unprecedented numbers recorded in the early 1990s. An interesting detail is the rise in the 1998 numbers, which may be the first indication of a shift, compared with the period 1994-1997.

[4] The exact timing of the recession can be debated. The general opinion is that the major shift was in 1989/90.

[5] In 1991 banks in Norway made provisions for credit losses amounting to NOK 20 100 million, compared with NOK 4 500 million in 1987

Year	Bankruptcies; limited liability	Bankruptcies; Total*
1990	Not registered	3 814
1991	Not registered	4 926
1992	Not registered	5 749
1993	Not registered	5 158
1994	2 224	3 664
1995	2 195	3 500
1996	2 141	3 458
1997	2 054	3 333
1998**	2 493	3 472

Table 1: Number of Bankruptcies in Norway, numbers for limited liability companies and the total number of bankruptcies registered
Source: Statistics Norway (Bank og Kredittstatistikk 2/96 and 2/98)
*) These figures include personal bankruptcies. Registration method changed fin 1994
**) Source: Dun & Bradstreet Norway.

Norway's bankruptcy legislation states that a debtor shall begin bankruptcy proceedings if the debtor is insolvent and a bankruptcy petition is submitted. Either the debtor or any of the debtor's creditors can submit a petition for bankruptcy. *The debtor is considered insolvent if he is unable to fulfil his economic obligations as they mature. He is not to be considered as insolvent if his property and income are sufficient to cover the obligations given time to be liquidated.* The Norwegian penal code §283a require a debtor to petition for bankruptcy when the debtor has reason to believe that the business is run at the expense of the creditors.

How to interpret the bankruptcy legislation is not totally obvious. There is a discussion whether a creditor can petition for bankruptcy if bankruptcy appears to be an inevitable outcome, but is not currently present[6]. There is currently no agreement on this issue. A related discussion is how to resolve a petition for bankruptcy if a company is technically insolvent, but has the prospect of future profits to cover financial obligations.

Bankruptcies result in losses to the affected creditors. Langli (1994) investigated 192 randomly selected bankruptcy petitions in Oslo from 1992. He finds that the creditors, in total, received only NOK 71.6 million of total secured claims of NOK 708 million. This is a 90% loss of secured claims, and an average loss of NOK 3.3 million NOK[7] per bankruptcy petition.

[6] Karnov (1996) presents the arguments used in this discussion.

[7] Figures are in nominal values.

Introduction

NOU (1993) reports that the city courts in Oslo charge an advance payment of NOK 100 000 at the start of debt settlement proceedings. A result of this is that too many debtors continue operations too long. The result is a hopeless debt settlement in the start of the bankruptcy proceedings.

To my knowledge, only Bardos (1998), a French study, has used more than 10,000 observations to investigate credit risk. For Norwegian data, this study is unique. Access to the complete population of limited liability companies permits an investigation of the entire business sector. The goal of the thesis is to describe the credit risk facing the Norwegian banking sector. To achieve this goal, bankruptcies, as a proxy to assess the risk of financial distress and default on loans, is used.

This thesis continues as follows: The next chapter is a brief summary of literature relevant to this thesis. Chapter 3 discusses the methodology most commonly used in credit risk measurement and chapter 4 describes the data set used. Chapter 5 presents both the models for risk analysis used in the Central Bank and the proposed model and Chapter 6 provides concluding remarks.

2 LITERATURE

In the development of financial distress models, it has not been much focus on economic theory. Two strands of literature explain why businesses disappear from the economy. A branch of industrial organisation theory investigates company exit, whilst financial analysis investigates financial distress. A complete survey of both these fields is beyond the scope of this thesis. Both fields provide models describing the mechanisms of exit or financial distress. For the data available in this study, many of these models are of little use as they depend on information that was not available to the author. This chapter introduces studies that are deemed relevant. The first and second sections introduce theoretical studies and the third section introduces empirical studies.

2.1 Industrial organisation theory

I this thesis, I have focused on the branch of industrial organisation theory that explains company exit. A common interpretation of company exit is company shutdown due to bankruptcy, mergers and acquisitions, voluntary liquidation or compulsory liquidation. To my knowledge, there are no comprehensive surveys over exit literature[8], but Hart (1995), Martin (1993) and Tirole (1989) report some of the different findings and models. In addition, there is special issue of the International Journal of Industrial Organisation focuses on plant turnover and growth patterns[9]. Martin (1993) states that exit is linked to entry with a revolving door. Therefore, an investigation should focus at only one of the two.

Jovanovic (1982) formulates the selection model. It describes why small companies can grow faster than large companies, and why small companies are less likely to survive. The hypothesis is that a company is in a continuous process of learning. Assuming company cost to be randomly and normally distributed among companies, the distribution of true costs in the economy is known to all, but no company knows its own true cost. With time, a company will observe costs that are normally distributed around the true cost of the company, i.e. with time, the company gains more knowledge about the true costs. Young companies have less

[8] Much work in industrial organisation theory has been done on entry and entry deterrence, rather than exit.

[9] International Journal of Industrial Organization, Vol.13, No.4 1995

information about their true cost and will be a more heterogeneous group as they base investments on less information. Cost observations below true costs induce excessive investments, this happens as management overestimates the prospects for future profits. A company will decide to exit if evidence arise that the true cost is too high. As time passes, company growth will adjust according to true costs. Successful companies will gradually reduce growth, adjusting operations to the correct true costs. Unsuccessful companies will be recognised by that they either grow too quickly, or remain small until exit[10]. The lack of information among young firms suggests a higher rate of exit for young companies compared with older companies.

Klette and Mathiassen (1996) find support for the selection model. They use both age and a variable for productivity to investigate company exit. They find that age is significant in explaining exit even when adjusting the system for productivity. They indicate that age contains information not covered in the productivity variable they use.

Another theory used to explain company exit is the vintage theory of Johansen (1959). He proposes that the age of capital equipment in a manufacturing company affect the decision for exit. Entrants employ more capital equipment with new technology than older companies. This reduces the competitiveness of older capital equipment, which in turn leads to the exit of older companies. The result of this is an observable, continuous update of the capital equipment within an industry. The vintage capital theory suggests that exit rates will increase after a company reaches a certain age. Salvanes and Tveterås (1998) use both the age of plant and the age of capital equipment as variables for explaining exit. To investigate the selection model they use plant age is used and to investigate the vintage capital model they use the age of capital equipment. They find support for both models, but the support differs among industries.

The vintage capital theory is not likely to hold for distress modelling. The reason for this is that the theory describes effects at plant level that need not hold at company level. It is reasonable to believe that it is easier for an already established company to open a new plant

[10] Extreme observations have a larger impact on young companies than on older companies as young companies have fewer observations of costs on which to base their decision..

than for a new company to enter[11]. Plant turnover therefore seems less important for distress prediction.

The idea of vintage technology has been extended to explain the updating of technology or a cleansing of productivity during recessions. Articles like Caballero and Hammour (1994, 1996)[12],[13] and Aghion and Howitt (1992) describe shakeout effects/creative destruction. Based on the vintage capital models they explain counter-cyclical exit rates. When the economy is in a boom, inefficient companies survive due to the slack and optimism in the economy. In addition, the competitive pressure on new entrants is small. This results in an accumulation of inefficient companies in the economy. When the economy enters a recession, a demand shock will lead to exit of inefficient companies. After the recession, the economy has increased productivity due to the cleansing.

Ghemawat and Nalebuff (1985) develop a model for exit in a declining market with oligopolistic competition. Extensions are discussed in Martin (1993). Ghemawat and Nalebuff (1985) find that in a declining industry where companies have a fixed level of production[14] the largest companies will exit first. The reason is the possibility for monopoly power. When facing identically sized markets, the smaller company will experience higher monopoly profit per unit then the larger company. In addition, the smaller company is able to keep up production longer into the time horizon than the larger company[15]. This knowledge is available to both companies and makes the large company exit first[16]. The authors make some extensions to the model. In an industry composed of multiple companies of fixed production, the argument above will still hold. The largest company exits first. Benefits from economies

[11] Older companies have established products in the markets and have therefore more credibility in the capital markets.

[12] These articles examine the labour market; job destruction is generally more responsive to the business cycle than job creation. This implies that company exit is more responsive to the business cycle than company entry.

[13] They report of similar articles and empirical findings in Breshnahan and Raff (1991,1992), Blanchard and Diamond (1990), Grossman and Helpman(1991) and others, but these articles were unavailable to the author.

[14] Examples they use are mostly from the refining and chemicals industry. The start-up costs for a steel mill are very high compared with the costs of keeping it running. In many cases, it is also necessary to operate close to capacity for the chemical processes to work smoothly.

[15] Declining demand will eventually force any producer out of the market. Unable to change the level of production the larger company will experience over-capacity earlier than the smaller company.

[16] The large company can not make credible threats to make the smaller company exit first. The small company has a larger profit potential as a monopoly compared with the large company.

of scale must be very large for the large company to be able to make threats credible. A last case is variable capacity; the large company will then start reducing capacity first.

Theories of incomplete contracts have a quite different approach to exit. Hart (1995) treats incomplete contract models that link design of capital structure to the liquidation of a company. Capitalisation of a company can be made in different ways. Incomplete contracts use the idea that any contract distributes bargaining power between agents. The design of the contract determines the operation of a company. Incomplete contract theory assumes that there exists an optimal timing of bankruptcy. The different agents involved in the company have different incentives and the capital structure is essential in questions of dissolution or bankruptcy. Generally, a too high debt will lead to dissolution when the company should maintain production and a too high equity share will result in maintaining production when the company should be dissolved. The reason is that debt restricts management too much and equity lacks control over management.

Industrial Organisation theory has also focused on the possibility that one company has the means to induce exit of other companies. Proposed theories are predatory pricing, product proliferation, hostile take-overs and more. The theories differ significantly and have different implications for the economy. They often focus on industry-specific or company specific information. I find that these theories either lack relevance to this study or that they depend on either micro-level information or market-specific information that is not available for this investigation.

The models treated above provide different insights into the analysis of exit. The selection model has two important implications: young companies will have a higher rate of exit than established companies and young companies are more heterogeneous. The vintage capital model implies that older companies will tend to exit with the introduction of new technology. The theories combined predict a U-shaped relationship between age and exit. Either of these two theories can be examined empirically by the use of the age of the company or the age of technology implemented in the capital equipment.

The exit model developed by Ghemawat and Nalebuff (1985) predicts that larger companies will tend to exit first when the market is declining. The incomplete contract theory predicts that, ceteris paribus, companies with an excessively high equity-loan ratio tend to live too long and companies with a too low equity-loan ratio will tend to exit too early. This is difficult to implement empirically, and expected to add heterogeneity in the bankruptcy process.

2.2 Financial statement analysis

Beaver (1966) and Altman (1968) undertook the pioneering work on financial distress analysis using financial data. Beaver (1966) uses univariate discriminant analysis and Altman (1968) uses multivariate discriminant analysis. Altman and Saunders (1998) describe the development of models that measure credit risk over the last 20 years. Surveys on financial distress literature using US data are Scott (1981), Zavgren (1983), Jones (1987), Keasey and Watson (1991). A survey on empirical work for non-US data is Altman and Narayanan (1997). A survey in Norwegian is Bruflot (1993). Olsen (1991) includes a comparison of empirical studies of Norwegian data.

A cash-flow model motivates Beaver (1966). The idea is that a company is a reservoir of liquid assets. This reservoir is supplied by inflows and drained by outflows based on profit/losses from operations. If outflows are consistently larger than inflows, the reservoir will empty and the company will experience financial distress. Beaver (1966) assumes four effects in his study:

- The larger the reservoir, the smaller the probability of failure
- The larger the liquid-asset net inflow from operations, the smaller the probability of failure
- The larger the amount of debt, the greater the probability of failure
- The larger the expenditures for operations, the greater the probability of failure

The model lacks a technical formulation and has no proposals to weighting for the different effects. The theoretical contribution from this model is limited to the fact that a company goes bankrupt when it runs out of liquid assets (Keasey and Watson (1991).

Another attempt to model the development into bankruptcy is the use of the gambler’s ruin model. Feller (1968) developed the gambler’s ruin model in probability theory. Gambler's ruin is a model where a gambler wins or loses money by chance. The gambler starts out with a positive, arbitrary, amount of money. The gambler wins a dollar with probability p and loses a dollar with probability (1-p) in each period. The game continues until the gambler runs out of money. Gambler's ruin is a mechanical description of the gambler's wealth.

The first proposal to implement the gambler's ruin model for financial distress prediction is Wilcox (1971), and is empirically tested in Wilcox (1976). The assumptions are that a company starts with an amount, *K*, of capital/assets, and Z indicates the periodical change in K. *Z* is a random number that occurs with a positive probability. It can be either positive or negative and it describes the cash flows from operations. Liquidating assets is the only method to cover negative flows. A company goes bankrupt in the next period if $K+Z<0$[17]. By modelling the flows from operations, it is possible to find expected time of distress.

Scott (1981) states that there are two major difficulties with the gambler's ruin model when predicting bankruptcy. First, the company has no access to the securities markets. Second, the cash flows are results of independent trials and managerial action cannot affect the results. Scott (1981) attempts to extend the gambler's ruin by introducing an imperfect access to external capital. I do not consider that this extension adds significantly to the original, apart from adding costs of external capital when finding a bankruptcy criterion.

Argenti (1976) discusses causes for bankruptcies. His states that most companies that go bankrupt follow one of three general patterns. These patterns are recognised by the following:

- *Companies that are unsuccessful:* unable to establish a foothold in the market, the companies go bankrupt after a reasonably short time. Most companies in this category go bankrupt within the first 5 years of establishment. The company is based on too lofty business ideas, and it is often characterised by weaknesses in management and budgetary control.

- *Companies that are too successful:* A rocket-start with a high demand for their products makes these companies grow quickly. The normal way of financing of the expansion is through loans. The organisational structure and control routines are unable to keep up with

[17] When formulating an empirical model the following criteria is used:

$\frac{Z-\mu_Z}{\sigma_Z} < \frac{K-\mu_Z}{\sigma_Z}$ where K is the liquidation value and μ and σ is the means and standard deviation used to normalise the expression.

the rapid expansion. Expenses become uncontrolled and lead to bankruptcy. This type of company normally goes bankrupt in 4 to 8 years.

- *Companies that stagnate:* older companies that have been operating in the market for decades. The company does not react to changes in demand and competition and go bankrupt.

2.3 Empirical literature

Empirical investigations of financial distress generally lack a theoretical foundation. While not regarded as a problem in the field, it has led to a lack of consensus on independent variables. Much of the research made is in private credit institutions, and the research results are not generally available. Of available studies, the two most commonly used empirical methods are multivariate discriminate analysis and discrete dependent variable models like probit and logit[18]. Chapter 3 presents these methods.

The majority of the financial distress prediction studies made are on US data. The univariate analysis by Beaver (1966) is considered the first empirical study for predicting financial distress. Altman (1968) use multivariate discriminant analysis and develops the Z-score model, a model where a company is given a credit score based on 6 financial ratios. He uses 33 bankrupt and 33 non-bankrupt companies. Altman *et al.* (1977) extend the analysis of Altman (1968) and introduce the ZETA[19] model. It uses multivariate discriminant analysis to identify 7 independent variables. These are:

- return on assets,

[18] Two only recently implemented methods the recursive partitioning algorithm and Neural networks. The Recursive Partitioning Algorithm (RPA) is a non-parametric, complex and computer intensive sorting algorithm. Based on explanatory variables RPA develop a sorting tree. It can be thought of as a multi-step DA sorting. A detailed description is given in Breiman *et al.* (1984). Altman *et al.* (1988) finds that RPA can outperform DA when predicting bankruptcy. Furthermore, they state that the results from RPA are difficult to interpret. RPA should therefore be used in together with other techniques.

Neural networks (NN) are also a computer intensive method. NN models are based on computer systems that develop “artificial intelligence”. Learning through experience gives the system an increasing ability to predict through recognising patterns. Altman *et al.* (1994) and Bardos & Zhu (1997) are studies investigating the powers of NN to predict financial distress. Altman *et al.* (1994) state that NN is potentially a powerful tool, but it is prone to several problems. NN systems have a tendency to over-fit and they are not easily interpretable. They conclude that NN should be used together with other models.

[19] To my knowledge, this is not an acronym, but the trademark name for a model owned by Wood, Struthers and Windthrop. (Altman *et al.*(1977))

- stability of earnings,
- debt service,
- cumulative profitability,
- liquidity,
- capitalisation,
- Size.

The model designed is prediction of financial distress up to 5 years before the event. Scott (1981) compares studies from the 1970s and finds the ZETA model the most convincing. The reason for this is that the ZETA has both high discriminating properties and a low level of model complexity.

Ohlson (1980) is the first general financial distress study using logit analysis. In addition, it is also the first study using a representative population sample; he uses 105 bankrupt companies and 2,058 non-bankrupt companies. Ohlson (1980) states that the predictive power of logit seems to be lower than the previous studies using multivariate discriminant analysis[20]. A study comparing different financial ratios is Chen and Shimerda (1981). They compare 65 financial ratios used in 26 studies. They find that the different definitions convey mostly the same information. While not finding a rule for selecting variables, they conclude that information on the following seven indicators should be used: Return on investment, financial leverage, capital turnover, short-term liquidity, cash position, inventory turnover and receivables turnover.

Platt and Platt (1990,1991) propose to use industry relative financial ratios. The argument is that if ratios consistently vary among different industries, industry-relative ratios should contain extra information. They will also have two desirable properties: 1) companies become more comparable across industries and 2) the coefficients will be more stable over time. Platt and Platt prefer industry-relative ratios rather than unadjusted ratios[21]. One criticism is that the determination of industry composition is not obvious. «Should only domestic companies

---

[20] One reason for this can be the use of a representative sample.

be used or should companies in the world market be included, and if the world market is included, to what degree? Should observations be weighted when composing the ratio?» These are highly relevant questions, particularly for small open economies.

Boyes *et al.* (1989) states that regular credit scoring models are too narrow. The argument is that credit lenders are interested in profit maximisation rather than simple default classification. Regular classification does therefore not contain enough information for the bank. A continuous description of risk is more useful as it can used to determine expected earnings from individual groups. A bank can then adjust the credit granting decisions according to profit targets and different risk profiles.

Banque de France model for risk analysis

Bardos (1998) describes a system of business sector monitoring used in Banque de France (BF - the French central bank). BF employs a database of 1.6 million companies to assess individual companies. In addition to this, they have a detailed credit-scoring model called BDFI. BDFI is used to monitor the 160.000 companies that account for 90% of credit in the French banking sector.

Credit granting decision in not the purpose for the BDFI, rather it describes the risk of distress in the economy. BDFI is a linear discriminant analysis model with equal costs of misclassification. The system has seven risk categories. Each category has a predefined probability of failure. Encompassing both the risk of the individual groups, and an identification of group membership, BDFI is used for multiple tasks. The primary task is to ensure stability in the banking system, both by monitoring the aggregate development and the portfolio of individual banks. A secondary task is to analyse individual companies. The system describes not only the current risk, but also tracks how observations move between groups in previous years.

Norges Bank (the Central Bank of Norway) apply a similar model to the one in the French central bank. Bardos (1998) reports that similar models exist in other central banks in Europe, including the German Bundesbank, the Italian Centrale dei Bilanci, the Bank of Austria and the Bank of England.

---

[21] Industry averages are found by using statistics from the US. Internal Revenue Service.

Work on Norwegian Data

Few studies are available that use Norwegian data to predict financial distress. In recent years, Norges Forskningsråd (Norwegian Council for Research) has been funding a project investigating economic crime. Part of this project covered financial distress, but not in the same manner as in this article.

The Central Bank of Norway has developed a model to describe risk in the Norwegian business sector. A summary of aspects relating to the NB model of business risk is SND (1995). Eklund (1988) use multivariate discriminant analysis to evaluate the NB model. He sorts companies into three categories: distressed, indeterminate, and non-distressed based on the score from the multivariate discriminant analysis. Looking at multiple different combinations of variables, Eklund (1988) concludes that working capital and stock are not suitable variables, while retained earnings and self-financing are found useful for prediction.

Andersen and Halvorsen (1992) evaluate and test new specifications of Norges Bank model using logit analysis[22]. Evaluating the original model using a representative sample, they find that the model capable of correctly classifying non-distressed companies, while it only correctly classifies 8.6% of the distressed companies[23]. They find that the ratios *equity / total debt* and *cash flow / long-term debt* significantly explain bankruptcies, and the ratio *working capital /earnings from operations* does not.

Other Norwegian studies on financial distress prediction include Gjesdal (1995). He uses non-financial indicators to predict bankruptcy. He examines 254 limited-liability companies that went bankrupt from 9 April to 11 June 1994 and he uses three indicators:

- Failed to file the 1992 annual report with the Register of Business Enterprises (RBE)
- Resignation of auditor in the period 15 September 1993-20 March 1994.
- If any liens were placed on the company through legal action before 15 September 1993.

[22] The paper uses both representative samples and balanced samples. Much varying results give reason to strongly believe in a significant bias when using a balanced sample.

[23] This is based on a p=0.5 cut off point

The sample is compared with a control group of random companies. He reports that the resignation of auditor is registered in 42.3% of the bankruptcy cases. Furthermore, 26.8% did not file their annual report at the RBE. Both variables are significantly different from the observations in the control group.

A study focusing on the timing of bankruptcies is Knivsflå (1997). He uses a data set with 13,166 bankruptcies from 1991-1996 and describes the development of financial ratios for up to 6 years before bankruptcy. He investigates measures for profitability, liquidity and financial strength. Knivsflå (1997) finds that all ratios show a monotonous deterioration. Interestingly, he finds that smaller than average companies generally have lower ratios and experience a deterioration that is more severe than larger ones. Langli (1994) investigates the relationship between creditor losses and economic crime in bankruptcy cases. He finds that there is a significant relationship indicating that the presence of economic crime gives creditors higher losses.

An industrial organisation-study is Klette and Mathiassen (1996). They use 83,237 observations from 16,689 industrial plants to explain plant exit. Examining all companies that exit an industry, they find that age is significant for the exit decision. They report a U-shaped connection, where young and old companies have higher probability of exit. While the hypothesis of Jovanovic (1982) is indicated to be correct, support for the capital vintage theory is not found. Another industrial organisation study is Salvanes and Tveterås (1998) who also investigate plant exit. They find evidence for the existence of both a capital vintage effect and a shakeout effect during recessions. When testing for both effects simultaneously they are unable to find significant evidence for higher exit rates among old vintages of capital during recessions.

A critical remark

In much of the reported research, especially in the discriminant analysis framework, the data samples are subject to some critique. The first problem is the use of balanced samples[24]. The overrepresentation of financially distressed companies in a sample gives reason to believe that the results will not hold in general. The second problem is the way the non-distressed sample is selected. A normal method is to pair distressed observations with non-distressed observations so that the observations are of similar companies. The third problem is how observations are weighted. Normally, the researchers weights observations by using different costs on the two types of misclassification errors[25], but rarely base the decision of misclassification costs to empirical findings.

Zmijewski (1985) investigates two more issues, the «choice based sample» bias and the «sample selection» bias. The first is related to the issue above-mentioned problem, an extremely low frequency rate of distressed companies. This has the effect of generating a non-random sample, and the validity of the results are suspect.

The ‘‘sample selection’’ bias arises when a non-random sample from the population is used. An example of this is investigation of loan default. The common sample to use is a population of granted loans, ignoring all the rejected applicants. The difference between the two biases is then that the “choice based sample ” bias is related to how the sample is composed internally and the “sample selection” bias is related to any connection between available data and the phenomenon investigated. There is a possibility that the data used in this thesis is subject to the sample selection bias.

Zmijevski (1984) investigates the effects of the two biases related to predicting financial distress. He finds that the first bias is significant when using a skewed ratio of observations. He also reports that the second bias exists, but does not seem to affect the statistical inferences or overall classification rates. Boyes *et al.* (1989) find that results can be improved when incorporating the credit granting decision into a model explaining the risk of loan default[26]. Eisenbeis (1977) discusses the problems related to financial ratios. He states that financial ratios are more often not-normally distributed than they are normally distributed. The following chapter discusses this further.

---

[24] A balanced sample has an equal proportion of distressed companies and non-distressed companies.

[25] The type of misclassification errors here is either classifying a company as bankrupt when it actually survives, or classifying a company as a survivor when it actually go bankrupt. For a bank, it is reasonable to believe that it has different losses to the different types of wrong predictions. The goal of a prediction model is normally to minimise the cost of misclassification

[26] Boyes *et al.*(1989) investigate profit-maximising behaviour in the credit granting process of banks. They use bivariate probit to model both the credit granting decision and loan default.

3 METHODOLOGY

Several approaches have been used to find determinants of financial distress. A dominating strand of research is discriminant analysis. Discrete response models mostly in the form of logit and probit are the most commonly used alternative methods. In this chapter the different methods are presented and discussed.

3.1 Discriminant analysis

Discriminant analysis is the most common method for separating companies that will experience financial distress. The idea is to classify companies into groups based on one or more variables. Discriminant analysis is a combination of finding the best vector of explanatory variables and splitting a population into two sub-populations based in predefined population characteristics. Discriminant analysis has been applied in most scientific fields. In distress prediction, it is normal to consider a distress observation as a success and a survivor observation as an unsuccessful.

Beaver (1967) uses univariate discriminant analysis. This is a simple sorting rule where the value of one variable, x, is used to separate observations into dichotomous categories. Discriminant analysis specifies a cut-off point, x^* where $x<x^*$ places an observation into category 1, and category 2 otherwise.

As with univariate discriminant analysis, multivariate discriminant analysis is a method for placing an individual into one of two sub-populations. A vector of characteristics, $x=(x_1, x_2, \ldots, x_n)$, is used. The vector enters as elements in a linear value-function. This value function is then used as the discriminating variable. Fisher (1936) initially suggested the method:

$$D(X) = \mathrm{X}\hat{a} \tag{3.1}$$

where X is a vector of x_i and β is the vector of regressor coefficients The coefficients vector $\boldsymbol{\beta}$ is chosen to maximise the ratio of the squared difference between the means of two groups, μ_1 and μ_2 , divided by the variance of **D**. Technical formulation; μ_i is the mean of x_j, and Σ is the variance-covariance matrix for **X** , where $\Sigma = \Sigma_1 = \Sigma_2$, then the variance of D(X) is $\boldsymbol{\beta}'\Sigma\boldsymbol{\beta}$. The Max-problem is then:

$$\underset{\hat{a}}{max} \frac{(\hat{a}'(\mu_1 - \mu_2))^2}{\hat{a}\,\text{Ớ}\,\hat{a}} \tag{3.2}$$

When the coefficients have been estimated, (3.1) will have a functional value that can be used to separate observations into one of the sub-populations:

$$\begin{aligned} &\mathrm{D(X)} < \mathrm{k} \Rightarrow \{\text{Successful category. (i.e. distressed)} \\ &\mathrm{D(X)} \geq \mathrm{k} \Rightarrow \{\text{Unsucessful category (i.e.non - distressed)} \end{aligned} \tag{3.3}$$

The ordinary cut-off value for the discriminating function is to use k=0. Before and during the estimation process there are some issues to be considered. The analysis brings about two possible errors: Classifying an observation as successful, which turns out to be unsuccessful, and classifying an observation as unsuccessful when it is successful. For many phenomena, it makes a difference which types of error occur. To make the estimated model sensitive to this it is vital to attach costs to the two types of error and include this in the estimation process. So in addition to (3.1) – (3.3), the following must be done:

- Predefine a cost for wrongly classifying a distressed company as non-distressed
- Predefine a cost for wrongly classifying a non-distressed company as distressed
- Find the cut-off value that minimises the sum of classification costs for the sample.

The linear discriminant analysis method depends on the following assumptions:

1. The distribution of **X** is multivariate normal
2. The variance-covariance matrices are equal
3. The prior probabilities for group membership are known
4. The means, μ_1 and μ_2, and the variance-covariance matrix are known.

Violation of assumption 1 makes the estimator inefficient and inconsistent. Violation of assumption 2 there is need for a quadratic formulation of the discriminating function. Violation of assumption 3 and/or 4 can be adjusted for using the data sample.

Problems attached to discriminant analysis

Gessner *et al.* (1988) report that one effect of discriminant analysis is that multicollinear variables can increase how the model fit. Negative correlation between the variables will always increase fit. Eisenbeis (1977) sums up most of the problems related to discriminant analysis when applied to economic and financial data in 7 sections. Three important issues are:

Distribution of the variables: Discriminant analysis assumes normally distributed variables. For economic and financial data, this is most often not the case. Some researchers assume that the deviation from normality has only minor effects on the results, but Eisenbeis reports findings showing that linear procedures are quite sensitive to deviation from non-multivariate normality.

Interpretation of the significance of individual variables: in contrast to ordinary linear regressions, discriminant analysis does not have unique coefficients. Each of the coefficients depends on which other coefficients are used in the estimation. There is therefore no way of determining the absolute value of any coefficient. Eisenbeis (1977) reports several methods for getting around this problem, but finds that they are not satisfactory. The proposed methods are usually based on equal group dispersions.

Choosing appropriate a priori probabilities and/or cost of misclassification: A major weakness of discriminant analysis is dependence on a relatively equal distribution of group membership. If one group of the population is larger than the other, discriminant analysis will ordinarily classify all observations in this group. The only method for solving this is to choose a priori probability for group membership. This method will be ad hoc, especially if there is reason to believe that the group membership probability changes over time. Classification models will only be correct for the same period that of the estimation sample.

3.2 Probit/logit models

Logit and probit are methods for explaining variables belonging to the exponential family. They are therefore able to handle a large group of variable distributions. While providing very similar results, the difference between the two is that logit has slightly thicker tails than probit. Several studies that develop both the logit and probit models exists. A brief study is Maddala (1983) and a more thorough treatment is Aldrich and Nelson (1984). A thorough treatment of maximum likelihood estimation is Greene (1993).

Both logit and probit are designed to model a discrete endogenous variable. The discrete variable is a reflection of an underlying continuous response variable. The continuous variable can be observed in either of two discrete intervals.

A technical formulation is: A continuous response variable y_i* can be explained by a linear regression, where **β'** is a vector of constants, and $\mathbf{X}_i$ is a vector of explanatory variables:

$$\begin{aligned} & y_i^* = \beta'\mathrm{X}_i - \varepsilon_i \\ & \wedge \ E(\varepsilon_i) = 0 \end{aligned} \qquad (3.4)$$

y_i* is not observable, but an aspect is observable through the dummy variable y_i. The dummy variable is defined by:

$$\begin{aligned} & y_i = 1 \text{ if } \mathrm{y}_\mathrm{i}^* > 0 \\ & y_i = 0 \text{ otherwise} \end{aligned} \qquad (3.\ 5)$$

It is important to note that y_i^* and y_i are different variables. As stated in (3.4), $\boldsymbol{\beta}'\mathbf{X}_i$ is not $E(y_i|x_i)$, but $E(y_i^*|x_i)$.

Using (3.4) and (3.5), together with probability theory, the following relations hold:

$$\Pr(y_i = 1) = \Pr(y_i^* > 0) = \Pr(\beta'\mathrm{X}_i - \varepsilon_i > 0) = \Pr(\varepsilon_i < \beta'\mathrm{X}_i) = F(\beta'\mathrm{X}_i) \qquad (3.6)$$

F(•) is the cumulative distribution function for the error term, ε_i.

The difference between logit and probit is a difference in the assumption of the error term. The logit model assumes a logistic distribution of the error term, and the probit model assumes a normal distribution of the error term. With identical logistically distributed error terms, the formulation of the logit model is:

$$F(\beta'\mathrm{X}) = \frac{\exp(\beta'\mathrm{X}_\mathrm{i})}{1 + \exp(\beta'\mathrm{X}_\mathrm{i})} \qquad (3.7)$$

With identical normally distributed error terms, the formulation of the probit model is:

$$F(\beta' X_i) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\beta' X_i} \exp(-\frac{t^2}{2\sigma^2})dt \tag{3.8}$$

A comment about the probit model is that σ and **β’X**$_i$ are impossible to separate in the estimation. This implies that the expression must be formulated with σ=1:

$$F(\beta' X_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta' X_i/\sigma} \left\{ \exp(-\frac{t^2}{2\sigma^2})dt \right\} \tag{3.9}$$

The logit and the probit models give very similar results. I decided to use the probit model, this because industrial organisation theory usually assumes normally distributed company types[27]. In addition, normality is a result from large sample theory.

Interpretation of the coefficients in a probit model is through the marginal effects, i.e. the partial derivatives of the model:

$$\begin{aligned} &\Pr(y_i = 1) = F(\beta' X_i) = \Phi(\beta' X_i) \\ &\frac{\partial \Pr(y_i = 1)}{\partial X} = \frac{\partial \Phi(\beta' X)}{\partial X} = \phi(\beta' X) \cdot \beta \end{aligned} \tag{3.10}$$

where ϕ(β’X) is the density function for the normal distribution. While the coefficient will have different effects on the system depending on the other variables, it can be seen that the sign of the coefficient gives the direction of the effect from the variable on the probability of outcome.

The likelihood function is then formulated with the observed y-values. The values of y are realisations of a binomial process with probabilities given by (3.6). In a sample with repeated trials, where x_i varies, the likelihood function (the maximum likelihood method is dealt with below) is then the product of probability adjusted realised outcomes, or:

$$L = \prod_i \left[F(\beta' X_i)\right]^{y_i} \left[1 - F(\beta' X_i)\right]^{(1-y_i)} \tag{3.11}$$

[27] One example is the selection model of Jovanovic(1982)

3.3 Maximum Likelihood Estimation

Maximum likelihood estimation is a method for finding an asymptotically efficient estimator for a set of parameters. This means that when a data-sample grows in size, the maximum likelihood estimator will approach statistical efficiency, i.e. the method depends on large data-samples to give good results. Pratt (1981) proves that the likelihood function is concave, i.e. only one maximum.

In this presentation, the variables are equivalent to the variables in the probit presentation. When investigating a sample of n observations, y_j, from the same distribution it is possible to formulate the joint density function for the sample. The joint density function is the product of the individual density functions assuming that the observations are i.i.d[28]. This joint density function, known as the likelihood function, defined as a function of an unknown parameter vector θ:

$$\begin{aligned} & f(y_1, y_2, \ldots y_n, \theta) = f(y_1, \theta) * f(y_2, \theta) * \ldots * f(y_n, \theta) \\ & = \prod_i f(y_i, \theta) = L(\theta \mid \mathrm{Y}) \\ & \text{where Y is the vector of observations } \mathrm{y_i} \end{aligned} \tag{3.12}$$

A logarithmic formulation of (3.12) is easier to use, The logarithmic formulation is known as the log-likelihood function. Because density functions by definition are non-decreasing and logarithms are monotonic transformations, the maximum of (3.12) can be found by maximising the log-likelihood function:

$$\begin{aligned} & \max_\theta \left(L(\theta \mid \mathrm{Y}) \right) = \max_\theta \left(\sum_i \ln\left[f(y_i, \theta) \right] \right) \\ & \text{Solution to the problem is given by: } \frac{\partial \ln \mathrm{L}(\theta \mid y_i)}{\partial \theta} = 0 \end{aligned} \tag{3.13}$$

[28] i.i.d. - Independent and identical distribution.

It means that every observation is a trial from the same density function. In addition, the outcome of one trial does not affect the outcome of any other trial.

In practice, the estimators are found through numerical approximation. It is therefore common to state both the unconstrained estimates of the log-likelihood function, i.e. the value with only a constant element, L_R, and the maximum estimates L_U. Normally reported in logs, L_R is the log of (3.12):

$$\ln L_R = n\left[\left(P \ln P\right) + (1-P)\ln(1-P)\right] \tag{3.14}$$

where P is the proportion of $y = 1$ and n is the total number of observations

The maximum likelihood estimator has the following properties, cf. Greene (1993):

- It is consistent.
- It is asymptotically normally distributed.
- It is asymptotically efficient.

For a more technical specification of the properties, the reader is referred to Greene (1993). He states that maximum likelihood estimators have the minimum variance achievable by a consistent estimator.

The likelihood ratio test

For large samples there are three asymptotically equivalent, commonly used test procedures for testing the null hypothesis that the estimated coefficients are equal[29] to 0. Greene (1993) states that the test procedures produce equivalent results on large samples. I have therefore focused only on the likelihood ratio test.

The likelihood ratio test is a comparison of the constrained and unconstrained values of the likelihood function. The constrained value is L_R given in (3.14) i.e. the value of (3.12) with all the coefficients equal to 0[30], and the unconstrained value is L_U, i.e. the value of (3.12) when applying the estimated coefficients.

[29] These are: Likelihood ratio test, Wald test and the LaGrange Multiplier (LM) test. They are all based on the Chi-square distribution.

[30] In the probit framework this is equivalent of the hypothesis that the coefficient vector **β=0, i**.e. running the estimation with only a constant element.

Let L_R be the constrained maximum likelihood, and let L_U be the unconstrained maximum likelihood, then the likelihood ratio is given by:

$$\lambda = \frac{L_R}{L_U} \text{ where } 0 < \mathrm{L}_R < L_U \Rightarrow \lambda \in [0,1] \tag{3.15}$$

If λ is low, one cannot reject the null hypothesis. Formally, the likelihood ratio test statistic is formulated by:

> ***Likelihood ratio (LR) test statistic:*** *Under regularity, the large sample distribution of LR=-2\ln λ is chi-squared, with degrees of freedom equal to the number of restrictions imposed.*

When testing model specification must $LR>\chi^2(df)$, where df is the number of degrees of freedom for the coefficients of the unconstrained model to be significantly different from 0[31].

Measuring goodness of fit

Reporting the fit of likelihood functions is not as straightforward as with least square estimation. Greene (1993) describes the likelihood ratio index (LRI) as the most commonly used measure of fit for likelihood estimation (also known as McFaddens R^2).

LRI is a comparison of the estimated likelihood function, L_U, and the value of the likelihood function L_U.

The proposed LRI is similar to the ordinary R^2:

$$LRI = 1 - \frac{\ln L_R}{\ln L_U} \tag{3.16}$$

The hypothetical interval for LRI is [0,1]. LRI can only approach 1, but never attain the value. It is commonly assumed that to obtain a "perfect fit" L_R must equal 1. L_R=1 is equivalent to $F_i(\bullet) = 1$ when y=1, and $F_i = 0$ when y=0. However, if this is the case, then F_i is not a meaningful probability density function.

[31] The null hypothesis can be another constraint than coefficient equal to zero.

In the literature, it the uncritical assumption is that an increasing LRI implies an increasing goodness of fit. Greene (1993) states that the properties of LRI entail that there is no natural interpretation of the numbers between 0 and 1. As a further comment on measuring fit for likelihood models, Green states that the naive[32] model never has zero fit.

As an alternative to LRI, one can look at the predictive ability of a model by investigating the distribution of predicted probabilities for the two discrete groups. One prediction rule is:

$$y_i = \begin{Bmatrix} 1 & \text{if } \mathrm{P} > \mathrm{P}^* \\ 0 & \text{otherwise} \end{Bmatrix} \tag{3.17}$$

A common cut-off point is $P^*=0.5$. This rule is not satisfactory when one group is much larger than the other, probit estimation will then rarely predict probabilities above 0.5 and the model will always under (or over) predict.

3.4 Comparison of logit/probit and discriminant analysis

If the standard assumptions for discriminant analysis hold, logit, probit and discriminant analysis will give equivalent results. Gessner *et al.* (1988) compare OLS, probit, logit, linear discriminant analysis and quadratic discriminant analysis for use on binary dependent variables. They report that the five statistical techniques provide equivalent results empirically under one or more of the following conditions:

- the data do not violate any of the underlying assumptions
- the group covariance matrices are unequal
- the predictor variables are collinear

Testing data with log-normally distributed predictor variables, Gessner *et al.* (1988) find that logit and probit outperform the other techniques[33].

[32] The naive model is simply assuming that there will be n*P successes in an n-population, where P is the portion of successes.

[33] Logit and probit produce very similar results

Lo (1986) states that the logit model is more robust than discriminant analysis. This is equivalent with logit being an unbiased estimator under a wider range of circumstances than discriminant analysis. He shows that logit analysis is appropriate for any distribution from the exponential family. He concludes by stating that decreased computational costs make logit a more optimal method compared with discriminant analysis.

4 DATA SELECTION

In recent years, the Central Bank of Norway has developed a database for the Norwegian business sector. The database is named SEBRA (System for Elektronisk Behandling av Regnskapsanalyse - Computerised System for Accounting Analysis). The initial purpose of SEBRA was to monitor the loan portfolio of Statens Nærings og Distriktsutviklingsfond (SND - Norwegian Industrial and Regional Development Fund). A description of this work can be found in SND (1995).

The database has been expanded to contain all available annual balance sheets for companies with compulsory registration requirements. By law, many companies are obliged to register their annual statements[34] with Foretaksregisteret (Register for Business Enterprises) in Brønnøysund. The balance sheets are converted to an electronic format with the help of Dun & Bradstreet Norway (DBN)[35]. In addition to the balance sheets, DBN registers all bankruptcies and compulsory dissolution reported in Norsk Lysningsblad[36]. Currently SEBRA has records back to 1988, with the exception of 1992. There are no records on bankruptcies before 1991[37], thereby making analyses with a two-year horizon impossible for 1988.

4.1 Size of the data set

In the main analysis, data for the period 1993-1998 are used. Data for the period 1988-1991 were available, but they were not investigated initially. There are two reasons for using only 1993-1998. First is the fact that the Norwegian government implemented a tax reform in 1992. Due to out of the ordinary accounting methods used for the transition in 1992, this year

---

[34] All companies of limited liability are required to send in their annual reports.

[35] Assuming all balance sheets are free of errors when arriving at Foretaksregisteret, two independent error sources are possible for the SEBRA database. The first error from the reading/punching of the data from the original paper forms to an electronic form. The second error come due to differing implementation of the data entries used by DBN and the format of SEBRA. SEBRA has a reduced number of balance sheet entries so entries are added when going from DBN to SEBRA. This process is known to have some flaws in treating incomplete balance sheets.

[36] Norsk lysningsblad is published weekly and it lists which companies that have been petitioned for bankruptcy. It is the common source for information on bankruptcy petitions in Norway.

[37] Government regulations demanded that prior to 1991 bankruptcies these were deleted from the DBN database.

has not been included in SEBRA. Economic effects of the reform are discussed below. Second, the sample size for 1988 is approximately 50% smaller than the later years, making the year incomplete.

In the empirical investigation, I include only limited liability companies based on International System for Industrial Classification (ISIC) codes in the range 01000-74999[38]. The reason for using only limited liability companies is that this is a relatively homogenous group when it comes to accounting standards. Companies with ISIC classification above 75000 were excluded as they are primarily regarded as public sector services in Norway.

In SEBRA, each observation is a collection of information from an individual company. The following information is in principle available for all observations:

- a unique 9-digit identification number
- a 5-digit ISIC number
- year of establishment
- a financial balance sheet
- year of bankruptcy (if applicable)
- year of voluntary liquidation (if applicable) (Only for the years after 1992)

Many observations lack one or more of the above entries. Exclusion of these companies from the data is limited. Most of the missing entries appear in the financial balance sheet group. The treatment of these companies is discussed below.

Another important issue is at what point in time significant information is available. Most Norwegian companies use the calendar year as the accounting year. The performance of a company during year *t* is summarised in the annual report. The deadline for registering the annual report is 1 July, year *t+1*. The complete set of observations is available approximately one month later. Information on new bankrupt companies is available continuously through "Norsk Lysningsblad”, which is published year round.

The figure below summarises the stream of information relevant to an annual balance sheet.

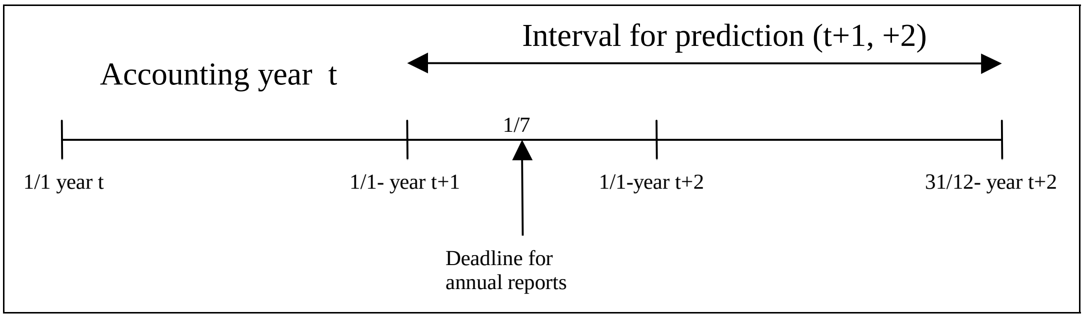

Figure 1: Timeline

The delay in publishing annual reports presents a problem for analysis as the data set lacks information on bankruptcy month. Causality that biases the analysis can therefore exist. The reason for causality is that a company that goes bankrupt during the beginning of year *t+1* might disregard its reporting obligations for annual reports from year *t*.

It is difficult to give exact measures of the quality of the data. From small companies, balance sheets are expected to have a lower degree of correctness in reporting[39]. Some observations have clear, illogical, or impossible balance sheet entries. These entries were removed from the sample. To find erroneous observations, the data were subjected to some logical tests. These were:

- Is the annual statement for a year later than a recorded bankruptcy?
- Is the year of establishment later than the year of annual statement?
- Are any of the debt entries negative?
- Is the value of the inventory negative?
- Is the entry for production expenditures negative?
- Is the entry for wages and pension expenditures negative?

[38] In addition, companies classified in the range 65000-69999 were left out. This is group J - Financial intermediation, and this group is not included in SEBRA.

[39] One example is liquid assets. 6 737 companies have liquid assets of zero and should technically be (close to) bankrupt. Only 14 of these are registered as bankrupt in the same, or in the following year.

- Is the entry for total assets negative?

Observations deleted following the application of these tests amounted to 1.2-1.6% annually[40] of this group. Within the excluded group, a maximum of 1.4%[41] was recorded bankrupt. These results give no reason to assume a correlation between deleted observations and bankruptcies. In addition, some companies have an unclear business profile or are otherwise unclassified with ISIC code 00000. This group amounted to 0.36% and was excluded from the analysis.

4.2 Effects of the 1992 tax reform

It is reasonable to assume that the 1992 tax reform created a structural break in accounting methods[42]. The reform had both temporary and permanent effects. Temporary effects make observations for the fiscal year 1992 inaccessible in SEBRA. The permanent effects are difficult to determine with certainty, as 1992 was a turning point in the Norwegian business cycle. A more thorough discussion on the effects of the tax reform can be found in NOS (1993) and Fjeld, Gaaseide and Stensrud(1994).

Two temporary effects are mentioned. The first is that the tax reform of 1992 included a change in corporate taxation, a change from a system of progressive taxes with a top 50% marginal rate with a system with a flat rate of 28%. To permit a smoother transition, permission of individual companies were permitted to shift taxable income from 1991 to 1992. Fjeld, Gaaseide and Stensrud (1994) report the effect of this as a positive difference of approximately NOK 70 billion between values accounted for and tax liability.

The second temporary effect was a large increase in corporate dividends between 1991 and 1992. The new tax system changed the way to account for equity, increasing the amount of equity available for dividends. This increase in dividends occurred although average annual profits decreased significantly.

[40] The year 1996 has the highest number, 1303 deleted observations, and 1995 is the lowest with 1127 deleted observations.

[41] Peak year was 1993, where 18 of the deleted observations went bankrupt within the following two years.

[42] The reform included deletion of entries from the required balance sheet. It is unclear how these entries were distributed on the remaining posts.

The most significant permanent effect of the tax reform was a change from a progressively increasing marginal tax to a flat rate of 28%.

Determining the complete impact of the tax reform is not possible. Still, one important issue needs mentioning; the tax reform had a definite effect on the way of accounting for equity. This illustrated by the fact that the average equity-to-assets ratio increased from 1991 to 1992, even though the average company experienced decline in profits and increased dividends.

4.3 Preliminary Examination of the data set

This section provides a brief presentation of tendencies in the data set. The period of 1993-1997 was one of economic growth and prosperity in Norway. Interest rates and inflation were low and stable. Unemployment decreased steadily and government surpluses increased. The period also has significantly lower levels of bankruptcies compared with the previous four years. The economy was consistently on the upside of the business cycle. All figures presented below are CPI adjusted values if appropriate. The index year used is 1995.

The data set consists of limited liability companies. This is a fairly heterogeneous group of observations. The initial choice was companies in ISIC range 00000-74999, excluding section J: Financial intermediation. In addition, three ISIC sections appear to have no observed bankruptcies in the period. Hence, the following three sections were excluded:

- Section A: Agriculture, hunting and forestry
- Section C: Mining and quarrying
- Section E: Electricity, gas and water supply

This leaves the majority of the business sector in the data set. Table 2 below shows the distribution of the population on the different ISIC categories. The total number of limited liability companies has been increasing in the period with a relatively stable distribution across sections.

ISIC Section	1989	1990	1991	1993	1994	1995	19996	1997
Fishing	1.69	1.55	1.47	1.53	1.57	1.51	1.44	1.39
Manufacture	15.00	13.33	12.97	13.20	13.47	12.79	12.44	12.07
Construction	8.62	8.96	8.90	9.16	9.18	9.37	9.37	9.37
Wholesale and retail tr.	33.52	33.27	33.24	34.26	34.10	33.85	33.67	32.85
Hotels and Restaurants	2.69	3.48	3.68	4.02	3.85	4.02	4.12	4.07
Transport and Com.	6.52	6.95	7.01	7.33	7.31	7.26	7.15	7.26
Real Estate & business act.	31.96	32.47	32.72	30.49	30.51	31.20	31.81	32.99
Total	100.00	100.01	99.99	99.99	99.99	100.00	100.00	100.00
# of Observations	**51 033**	**76 107**	**81 566**	**83 094**	**77 167**	**85 787**	**92 100**	**96 188**

Table 2: Distribution of observations subject to sector
(Figures are given in percent of total observations used in the analysis)

Two ISIC sections are significantly larger than the others, sections G and K, respectively “Wholesale and Retail trade and Real Estate, Renting and Business Activities". Together these two constitute for two thirds of the sample. The third largest group is D "Manufacturing", gradually declining to 12% of the sample.

Table 3 below gives the distribution of observations in respect of age. The year 1994 is an anomaly year compared with the other years. The year 1994 has a much lower number of observations for companies aged 0-1 year[43]. The reason for this discrepancy is unknown to the author, and I found no reasonable solution to adjust for this problem.

[43] If the low numbers were caused from a low number of entrants then there would be very few observations in the category for 1 and 2 years. This is not the case and the problem lies in the collection/registration of the 1994 data.

Age	1 993	Percent	1 994	Percent	1 995	Percent	1 996	Percent
0	5 607	7.11	884	1.16	5 462	6.64	6 756	7.63
1	7 259	9.21	3 333	4.38	7 307	8.89	7 437	8.40
2	6 898	8.75	6 589	8.65	6 308	7.67	7 022	7.93
3	6 901	8.75	9 338	12.26	5 731	6.97	5 999	6.77
4	6 780	8.60	6 408	8.41	5 649	6.87	5 481	6.19
5	5 771	7.32	6 335	8.32	5 723	6.96	5 399	6.09
6	5 649	7.16	5 401	7.09	5 745	6.99	5 560	6.28
7	4 617	5.86	5 362	7.04	4 956	6.03	5 604	6.33
8	3 479	4.41	4 381	5.75	4 931	6.00	4 756	5.37
9	2 437	3.09	3 324	4.36	4 069	4.95	4 827	5.45
10	1 941	2.46	2 301	3.02	3 085	3.75	3 968	4.48
11	1 694	2.15	1 879	2.47	2 159	2.63	3 008	3.40
12	1 636	2.08	1 633	2.14	1 778	2.16	2 115	2.39
13	1 273	1.61	1 569	2.06	1 532	1.86	1 715	1.94
14	1 063	1.35	1 241	1.63	1 475	1.79	1 494	1.69
15+	15 837	20.09	16 195	21.26	16 319	19.85	17 444	19.69
Sum	**78 842**	**100.00**	**76 173**	**100.00**	**82 229**	**100.00**	**88 585**	**100.00**

Table 3: Distribution of observations subject to age

The lack of observations of young companies represents a problem with the 1994 observations. Generally, this age category accounts for a large proportion of the total bankruptcy population. In the sample, typically the first two years contains some 25-30% of the number of bankrupt companies. Expect therefore that regressions run on 1994 data will prove to give different results compared with the other years.

Development in equity levels

One important size in a balance sheet is the amount of equity in the company. Equity is considered to serve as a cushion to ease temporary financial difficulties. The higher the equity the longer a company can survive in periods of hardship. In Table 4, one can see that the equity levels and differences in equity have been increasing in the period. In particular, 1997 showed a significant jump both in level and in heterogeneity.

This corresponds with the hypothesis of shakeout effects. A period of growth and stability should be accompanied by an increasing number of inefficient companies. While the deviations among companies have been increasing, there has been a decrease in the relative number of companies with negative, or zero equity. In 1993, 23% of the population had negative or zero equity. In 1997, this had fallen to 17%. In absolute terms, the number of companies with negative or zero equity has been steady at around 17 000.

Development in debt levels

The opposite of equity, the level of indebtedness, gives an indication the timeframe in which a company will experience problems in an economic downturn. High debt is a drain on liquidity; if revenues drop, liquidity suffers. Another problem with high debt is the vulnerability to negative shocks in the macro environment. This type of shock often leads to higher interest rates, and decreased liquidity in the market for debt issuance.

Short-term debt is common among companies. More than 95% of the sample have registered some level of short-term debt. In the period, the level is relatively stable at NOK 3.1 million, as shown in Table 4.

Long-term debt is less common than short-term debt with some 63% of the observation having some level of long-term debt[44]. This ratio drops to 54% if deferred taxes are not included. The investigation of long-term debt is done with a slight twist. There is a small tendency for a decreasing ratio of companies with long-term debt, so I will look only at the companies that have long-term debt. It then seems that there is a clear indication of increased heterogeneity in the sample, especially for 1997. See Table 4 below.

Year	Revenue reserve	Short-term Debt	Long-term debt	# Obs.
1989	2207.6	8518.1	9923.9	51 033
1990	1773.2	5230.6	8748	76 107
1991	1660.7	4744.8	8491.6	81 566
1993	2734.9	3836.8	6453.9	86 044
1994	3058.3	4121.9	6344.7	80 021
1995	3090.3	3826.7	6949.6	88 881
1996	3393.7	3830.2	6281.8	95 171
1997	4517.8	4180.8	7515.3	96 188

Table 4: Development of mean; debts and revenue reserves 1989-1997.
Figures are in NOK 1000, adjusted for CPI.

[44] Includes postponed taxes considered as long term debt.

5 REGRESSION RESULTS

5.1 The Central Bank of Norway's model for risk

The Central Bank of Norway currently uses a model for risk classification. The initial purpose of the model was to monitor the loan portfolio of Statens Nærings og Distrikstuviklingsfond (SND – Government fund for regional and structural development). The model structure was developed with multivariate discriminant analysis to predict loan default. Three financial ratios are used to classify companies into 18 risk groups:

- Self-financing: Operating result, adjusted for taxes and depreciation of capital as a share of long-term debt
- Solidity: Revenue reserve as a share of total capital
- Liquidity : Liquid assets subtracted by short-term debt, divided by operating revenues

For the self-financing and the solidity ratios, three intervals are used: nominator/denominator are: smaller than 0%, between 0% and 20%, and greater than 20%. The third indicator is only a dummy for good or bad liquidity.

The risk model of Norges Bank is used on the SEBRA data set. I re-estimated the model with probit to make the results comparable to an alternative specification. As the model depends only on discrete intervals, I use dummy variables as explanatory variables. To avoid over-specification observations with ratios greater than 20% do not have a separate dummy.

The original model is a sorting mechanism and the fundamental properties are assumed to hold[45] when reinterpreting the model in a probit framework. Probit analysis finds the probabilities for bankruptcy for each group. The following functional form describes the implemented model:

$$\Pr(\text{bankrupt in year } t+1, t+2) = F(X_t) \tag{5.1}$$

[45] The use of dummy variables gives all observations in one group the same properties in a probit model. I.e. they are assigned the same probability/regarded as identical.

Where F(•) is the probit function described in chapter 3 and X_t is a vector consisting of the following dummy variables:

- Negative self-financing - Dummy for having a negative self-financing/long-term debt ratio
- Mid-range self-financing- Dummy for having self-financing/debt ratio in the 0-20% interval
- Negative solidity- Dummy for having a negative revenue reserve/total debt ratio
- Mid-range solidity - Dummy for having revenue reserve/total debt ratio in the 0-20% interval
- Negative liquidity - Dummy for having a negative liquidity indicator

A result of the analysis is that all companies in one category receive the same probability for bankruptcy. Four years of annual reports for 1993-1996 were used in the analysis to predict bankruptcies for 1994-1998. One weakness of the model is the use of long-term debt. Less than 2/3 of the sample has long-term debt leading many observations to lack a ratio. The standard procedure in Norges Bank is to assume that these observations are on the bounds, i.e. either -100% or +100% depending on the nominator. Observations with a negative nominator but with a zero denominator have been assigned to the <0% group. Observations with a positive nominator, but zero denominator are placed in the group of >20% ratio (i.e. respective dummies both have value 0).

Regression results

The table below shows the results of the four estimations. One can see that the coefficients and the pseudo R^2 are relatively stable for the four estimates. The pseudo R^2 are the same as the Mc Faddens R^2 described in chapter 3.

Regression Results

Variable	1993	1994	1995	1996
Negative solidity	0,877	-965	0,965	0,959
	(26,96)	(23,11)	(22,31)	(26,124)
Mid-range solidity	0,458	0,489	0,461	0,451
	(13,54)	(11,20)	(12,43)	(11,734)
Negative self financing	0,274	0,296	0,229	0,218
	(10,81)	(9,27)	(7,60)	(7,46)
Mid-range self financing	-0,274	-0,286	-247	-0,18
	(-6,60)	(-6,17)	(-5,76)	(-4,46)
Negative liquidity	0,24	0,161	0,231	0,257
	(10,42)	(5,54)	(8,66)	(9,92)
Constant	-2,7	-2,89	-2,81	-2,87
	(-101,47)	(-83,66)	(-96,18)	(-96,40
pseudo R2	**0,1256**	**0,1353**	**0,1116**	**0,1376**
# obs.	**83959**	**77956**	**86659**	**93008**
logL(0)	**-8090,1**	**-4972,3**	**-5677,5**	**-6115,7**
Log(5)	**-7073,6**	**-4299,6**	**-5044,1**	**-5274,2**
Chi-Sq.(5)	**2033**	**1345,6**	**1266,7**	**1683,1**

Table 5: Predicting bankruptcy for year *t+1,t+2* using annual report from year t**,**
z-values are reported in parenthesis

The estimates can be used to generate risk predictions. Each of the estimation results is then used to generate risk measures for the whole sample. The picture that arises is that all observations are assigned a low bankruptcy probability. In addition, there seem not to be too much difference between those companies going bankrupt and the ones that survive. Table 6 to Table 9 sum up the distribution of the predictions. Observations that go bankrupt receive a slightly higher (approx. 2%) predicted risk than the ones that survive, but they also have a higher variation.

Feil! Ugyldig kobling.

Table 6: Mean of predicted risk for companies observed as non-bankrupt or bankrupt, using 1993 data to predict bankruptcy in year 1994 and 1995.
Standard deviation in parenthesis

Feil! Ugyldig kobling.

Table 7: Mean of predicted risk for companies observed as non-bankrupt or bankrupt, using 1994 data to predict bankruptcy in year 1995 and 1996.
Standard deviation in parenthesis

Feil! Ugyldig kobling.

Table 8: Mean of predicted risk for companies observed as non-bankrupt or bankrupt, using 1995 data to predict bankruptcy in year 1996 and 1997.
Standard deviation in parenthesis

Regression Results

Feil! Ugyldig kobling.

Table 9: Mean of predicted risk for companies observed as non-bankrupt or bankrupt, using 1996 data to predict bankruptcy in year 1997 and 1998.
Standard deviation in parenthesis

Comment on the results

One can see from the regression results that a negative revenue reserve seems to be the most important factor in determining bankruptcy risk. The ratio self-financing/debt results are puzzling. Observations with mid-range self-financing are assigned a negative coefficient, i.e. decreased bankruptcy risk. Since this coefficient is compared with companies having a ratio greater then 20%, the results state that companies with mid-range self-financing are in strict terms less risky than companies with a self-financing >20%. This result is quite odd and is probably due to the interval size[46].

There are two issues to be concerned with when discussing the NB model for risk. The first is the definition of the trigger values. While a trigger at 0% seems logical, the second at 20% seems more ad hoc and a redefinition of these limits might improve the results. The second is the use of long-term debt, compared with total debt. To some extent, companies without long-term debt probably become incorrectly analysed. The use of long-term debt entails that too many companies are placed in the best group (i.e. ratio of 100%), while a lower ratio would be more appropriate.

5.2 Alternative model specification

In order to improve the model of Norges Bank, I investigated different specification possibilities. The primary idea was to improve use of available information. Financial ratios are continuous variables with information from an observation. A common assumption is that important information disappears when a continuous variable receives a discrete specification. I therefore constructed an empirical model that employs continuous variables where possible.

The purpose of the model is risk description. Preferably, the model should therefore give a correct indication of the individual risks in a company. The model presented below is partly an extension of the risk model of Norges Bank and partly a new model specification. The main differences are use of more information from a balance sheet and the use of continuous variables, rather than discrete variables. These extensions produce a more desirable result at the expense of model simplicity.

Probit estimation is used to develop the model. It could be argued that panel data theory would be more appropriate given the structure of the data. What is investigated is the probability of exit. A normal panel consists of observations from one unit over time. In the data set, there is a continuous change of companies from year to year[47]. Only the observations that survive for the entire period would be included in a proper panel, but these are also the "uninteresting" observations in that they do not go bankrupt. The variables are described below. The remaining sections describe the results and test the model on different samples.

Primary investigations were made on the 1993 sample. Different variable specifications have been tested and the ones described below were found to give the most consistent results. A description of various specifications is found below. The 1994-96 samples were used for investigating the consistency of the chosen variables. In the model presented, information from two years is used to explain bankruptcy the following year. The decision to explain only bankruptcies for one year was made to extend the number of independent estimations that could be used[48].

In the construction of financial ratios, there are two items to consider. The first is the possibility of creating variables that describes almost the same aspects of a company. The combination of ratios used in the analysis is important. The presented model is a system of ratios that go together well; it was found that other variables did not add significant information to the system. The second is the need for at least two entries from a balance sheet when constructing a financial ratio. The entries describe aspects of one company, but there is

---

[46] If the interval is poorly defined does the coefficient not need to be interpretable.

[47] For each year there are approximately 8000 companies (equals to 10% of the sample) observed that are not present in the next.

[48] At the outset of the analysis bankruptcy, figures for 1998 were not available. It would then only be possible to run two estimations if bankruptcies for two years was to be explained.

no direct dependency between the entries used in a ratio[49]. It is therefore not necessary for the denominator in absolute numbers to be the largest value. In some cases, it may be that a denominator is very small while the nominator is very large. I illustrate this with an example; both models presented in this thesis use the variable *revenue reserves/debt*. The interpretation of this variable is that the higher this ratio is, the less risky is the company. The problem arises with companies that has very low debt. If year-end revenue reserves are NOK 10 000 (approx. US$ 1 300), but debt is only NOK 1 000, the ratio becomes 10[50]. A ratio of 10 should indicate low risk, but revenue reserves of NOK 10 000 is no safety buffer. Very low debt will make the ratio very high, even though revenue reserves are low, with zero debt the expression give no meaning. Interpretation of the variable is not obvious. This produces non-credible ratio values[51]. I call this the extreme ratio problem. The selected solution to this was to create boundaries for the variables where the ratio tends towards infinity. Below is a discussion of the implementation on the individual variables.

Solidity

As a measure of the solidity of a company, I used the ratio *"revenue reserves to total debt*". It is reasonable to believe that this should be negatively correlated with bankruptcies. As mentioned in the previous chapter, revenue reserves are a buffer for the company to fend off periods of decreasing revenues[52].

Adjustments in the solidity ratio

The solidity ratio was bounded to the interval [-1,5]. The lower limit is quite logical, as revenue reserves in principle can in absolute terms never be greater than the amount of debt., This interval includes 96-98%[53] of the sample, depending on the year examined. Debt is by

---

[49] A ratio using variables dependent of each other can not be assumed to give relevant information.

[50] It is important to keep in mind that the information comes from the annual report, and the annual report is in some instances subject to window dressing.

[51] I find that 3-5 % of the observations have non-credible ratio values (i.e. values far outside the expected interval).

[52] A worrisome fact is that approximately 17.000 limited liability companies in the sample have a negative revenue reserve. These companies are surviving solely on the whim of the creditor. If a downturn of the Norwegian economy lead to a liquidity-crunch, will these companies instantly have problems.

[53] The percentage changes from year to year.

definition zero or positive, and negative ratios are purely the result of a negative numerator. Two types of adjustments were made:

- Observations with zero debt were set to the limit 5 if revenue reserves <0
- Observation with non-zero debt, but a ratio outside the determined bounds were set at the closest limit value (i.e. -1 and 5)

Dummy variables were created to keep track of adjusted variables.

Cash flow

As a measure of cash flow status, I use the ratio *cash flow to total debt*. Cash flow is measured by taking revenues after tax, before write-offs and extraordinary income. This ratio is expected to have a negative correlation to bankruptcy risk.

Adjustments in the cash flow ratio

This ratio was decided bounded by the limits [-1,1], as these limits included more than 97% of the sample. Debt is by definition zero or positive and negative ratios are purely the result of a negative numerator. Two types of adjustments were made:

- Observations with zero debt were set to the limit 1 if revenue reserves > 0, and -1 if revenue reserves < 0
- Observations with non-zero debt, but a ratio outside the determined bounds were set at the closest limit value (i.e. -1 and 1)

Dummy variables were created to keep track of adjusted variables.

Debt burden

To measure both the individual effect of macroeconomic tendencies and the impact of the debt burden, I use the ratio debt burden to cash surplus. Debt burden has a separate entry in the balance sheet, while cash surplus is constructed:

$$0{,}72 * (\text{revenue from operations} - \text{wage costs} - \text{cost of production inputs} + \text{financial revenues}) \quad (5.2)$$

It is a rough measure of revenue from capital and revenue from financial assets. The variable is adjusted with 0.72, as this is 1 - tax rate on income from business. For an individual company, a high debt burden will make a liquidity problem arise more easily.

Adjustments in the debt burden ratio

In contrast to the other two ratios, the debt burden ratio requires consideration of another issue. Negative observations must be a result of a negative denominator. A negative denominator is worse than having a zero or positive denominator. These companies are in severe distress. A debt burden and negative cash surplus only result in liquidity problems.

The boundaries of the debt burden ratio was set at [0,2] as more than 97% of the observations lie within this interval, handling outlying observations as follows:

- Observations with a negative ratio were given the value 2, this was done as these companies are inherently worse off than any company with a positive cash surplus.

- Observations with a positive cash surplus, but a ratio of above 2 were adjusted to 2.

- Observations with zero cash-surplus were adjusted to 2[54].

Accordingly, two dummy variables were created.

Age

A combination of the vintage capital idea of Johansen (1959) and the selection model of Jovanovic (1982) suggest a U-shaped relationship between age and exit. Different variable specifications for implementing this were tested[55], ending up with two explanatory variables for age: a linear function of age and a squared function of age. To see why this can indicate a U-shaped relationship one can look at the derivatives[56].

---

[54] No observations have both a nominator and a denominator of zero

[55] Dummy implementation, logarithmic and square roots.

[56] The probit model is a power function of the explanatory variables. But to investigate the individual relationship one can look at a linear representation

If f(•) is the density function for the normal distribution, and **β'X** is the product of the vector of explanatory variables and vector of coefficients will the derivative of the probability function with respect to x_k in **X** will be:

$$\begin{aligned} &\frac{\partial P}{\partial x_k} = \mathrm{f}(\hat{\mathbf{a}}\,'\mathbf{X}) \cdot \frac{\partial \hat{a}'\mathbf{X}}{\partial x_k} \\ &\frac{\partial P}{\partial x_k} = 0 \Leftrightarrow \frac{\partial \hat{a}'\mathbf{X}}{\partial x_k} = 0 \end{aligned} \tag{5.3}$$

On the assumption that f(**β'X)** is kept constant. The partial derivative of age is then found by:

$$\begin{aligned} &\text{Let } \hat{\mathbf{a}}\,'\mathbf{X} = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot (age)^2 + \beta_3 x_3 + \ldots + \beta_n x_n, \text{ then} \\ &\frac{\partial P}{\partial age} = \mathrm{f}(\hat{\mathbf{a}}\,'\mathbf{X}) \cdot \frac{\partial \hat{a}'\mathbf{X}}{\partial age} = \mathrm{f}(\hat{\mathbf{a}}\,'\mathbf{X}) \cdot (\beta_1 + 2\beta_2 \cdot age) \quad \text{for age} \geq 0 \\ &\frac{d^2 P}{d age^2} = \mathrm{f}(\hat{\mathbf{a}}\,'\mathbf{X}) \cdot 2\beta_2 \end{aligned} \tag{5.4}$$

Significance of both of the coefficients, β_1 and β_2 indicates that there is some form of non-linear relationship. If they have opposing signs, and $|\beta_1|<|\beta_2|$, then the hypothesis of a U, or an inverted U, relationship cannot be rejected. The first order condition defines the min, or max, point, and the sign of β_2 give whether it has a U- or an inverted U-shape.

From the selection model, it follows that a company has a decreasing risk of exit as it ages. This is equivalent of $\beta_1<0$. Furthermore does the vintage capital theory propose that old production plants will lose competitiveness, as production equipment grows older. After some age, it is therefore assumed that exit risk will start to increase again, which is equivalent of $\beta_2>0$. It should be mentioned that during the regressions, the functional form used for squared age is $age^2/100$. This was done to scale up the coefficients.

Dummy for missing balance sheet

A company in severe financial troubles usually has a collection of problems. Gjesdal (1997) finds that the lack of an accountant is a good signal of increased bankruptcy risk. If a company lacks an accountant, the annual statement will not be accepted in the Register of Business Enterprises. A positive relationship is expected between missing balance sheets and bankruptcy.

Dummy for restaurants

Two notorious groups in Norway are restaurants and bars. These groups are considered to contain many improper elements. In the analysis, it was tested with dummy variables for restaurants or bars. Both are expected to explain bankruptcy positively.

Dummy for real estate, renting and business activities.

The ISIC category "Real estate, renting and business activities" is a very heterogeneous group. Companies in this category generally have high levels of long-term debt, low liquidity, but also low business risk. Many of these companies have counterparts in the bar/restaurant industry. For many bars/restaurants, it is normal to have a separate company owning the premises of operation and renting it out to the actual business. This is a method for securing real estate property, while the valueless company (i.e. the bar/restaurant) can periodically go bankrupt.

In the analysis, this dummy is expected to have a negative relation to bankruptcy risk.

A note on alternative specifications

The reader might question the lack of macroeconomic indicators in the analysis. All estimations are a cross-sectional analysis where all observations come from the same period. A macroeconomic indicator would in this setting not be separable from the constant element[57],[58]. The macro environment has partly been included in the use of individual interest burden.

Platt and Platt (1990,1991) proposed the use of industry-relative ratios. The idea is that there are differences between the different types of businesses and by adjusting observations according to the sector that they operate in one can add information to the system. It could not be seen that regressing with industry-relative ratios added information to the system. Due to this and the arbitrariness in grouping observations discussed earlier, this approach was not

[57] Longer data series open for future investigations to fit macro economic variables with the constant element from the regressions.

[58] A German Study, Blevins, Lehment and Sjøvoll (1998), indicate a strong relationship between macro economic indicators and the rate of bankruptcy for limited liability companies.

investigated further. Another method for testing for differences among industries is by separation. It was tried using both dummy variables for the sector and running regressions for every sector independently. The variable coefficients for the different sectors did not prove different at a 95% level. Apparently, industry-relative ratios do add complexity to the system without improving the information content.

Most empirical studies from financial literature also include some measure for liquidity or cash flow. The focus was on working capital as a share of operating income, but it was not significant with time. One explanation for this might be that liquidity is relatively unimportant far from the time of bankruptcy[59]. An additional explanatory variable tested was the return on capital. In the appendix, the regression results include both the return on capital and working capital.

Ghemavat & Nalebuff (1985) propose that company size have an effect on exit. Ohlson (1980) uses the logarithm of (debt/GDP) as a proxy for company size. This and the level of operating revenues were tried as a measure for size. The variable was not included in the final model as the results indicated that bankruptcy risk was positively correlated with size, i.e. the larger the company, the greater the risk. This result drastically oppose economic intuition, hence it was disregarded. Additional variables from industrial organisation theory, like the level of R&D and the degree of sunk costs in a sector, were not implemented due to the lack of information on these issues. Incomplete contract theory proposes that an optimal combination of equity to debt exists, but no conclusive specification[60] was found in this thesis.

5.3 Regression results

As stated earlier, probit models were estimated. Each year was estimated independently. Regression coefficients and z-values are reported together with log likelihood values and pseudo-R^2. Pseudo-R^2 is the same as McFaddens R^2 described in chapter 3.

[59] Liquidity is normally the trigger for a bankruptcy petition, and can not be expected to be significant some time prior to distress.

[60] $(\text{Equity/debt})^2$ was tried and the result was inconclusive. The coefficient was significant for 1993, but not for the later years.

Regression Results

STATA 5.0 was used for the regressions. In some instances, STATA found indications of causality in the dummy variables[61]. This type of causality is equivalent of the endogenous variable having the same value for one category of the exogenous dummy variable[62]. Inclusion of the dummy variable is then meaningless. Variables dropped due to causality are noted together with the regression results. The regression can be described by the following function:

$$\Pr(\text{bankruptcy year } t+2) = f(\mathbf{X_t}, \mathbf{Z_{t+1}}) \tag{5.5}$$

Where $F(\bullet)$ is the probit function, X_t is a vector of balance sheet information presented previously and Z_{t+1} is information on missing balance sheets year *t+1*. Observations where bankruptcy was registered in year *t+1* were not used as observations.

The dummy for solidity<-1 was dropped by STATA due to causality for the t=1993, 1994, 1996, and the dummy for missing age data was dropped due to causality for t=1994, 1995, 1996. The estimation results are presented on the following page. The marginal effects of the coefficients are given in the appendix.

[61] For more information on STATA, the reader is referred to STATA (1997).

[62] An example is that for all observations where the dummy variable equals 0, no bankruptcy is observed, but the opposite does not hold when the dummy equals 1.

Regression results for bankruptcies year t+2				
Variable	**1993**	**1994**	**1995**	**1996**
Missing report year t+1	1,524	1,688	1,526	1,462
	(49.8)	(35.5)	(46.0)	(43.3)
Revenue reserves	-0,359	-0,187	-0,310	-0,358
	(-13.4)	(-5.7)	(-10.2)	(-10.5)
Revenue reserves>5*	0,956	0,177	0,861	1,053
	(5.8)	(0.8)	(4.6)	(5.0)
Revenue reserves<-1*	XX	XX	0,488	XX
Self-financing	-0,508	-0,634	-0,459	-0,469
	(-8.5)	(-7.7)	(-7.1)	(-7.2)
Self-financing>1*	0,582	0,316	0,302	0,300
	(4.7)	(1.6)	(2.0)	(2.0)
Self-financing<-1*	-0,640	-0,843	-0,600	-0,577
	(-6.1)	(-5.1)	(-5.1)	(-4.87)
Debt burden	0,208	0,310	0,158	0,102
	(4.5)	(4.4)	(2.6)	(1.6)
Debt burden>2*	-0,666	-1,071	-0,871	-0,545
	(-6.7)	(-6.9)	(-4.4)	(-4.0)
Debt burden<0*	-0,490	-0,704	-0,343	-0,307
	(-4,6)	(-4,4)	(-2,6)	(-2,2)
Restaurant*	0,386	0,246	0,317	0,188
	(6,4)	(2,7)	(4,9)	(2,7)
Real estate*	-0,272	-0,352	-0,308	-0,393
	(-7,3)	(-6,3)	(-7,2)	(-8,7)
Age	-0,015	-0,020	-0,035	-0,024
	(-5.0)	(-4.7)	(-10.2)	(-7.4)
Square of age	0,012	0,011	0,028	0,021
	(3.3)	(2.2)	(7.3)	(5.6)
Age missing*	0,265	XX	XX	XX
	(0.6)			
Constant	-2,438	-2,846	-2,364	-2,482
	(-84.0)	(-57.0)	(-77.0)	(-75.3)
pseudo R2	**0,337**	**0,389**	**0,341**	**0,334**
# obs.	**78 565**	**72 843**	**82 132**	**88 163**
ln L(0)	**-6 212,8**	**-3 206,0**	**-5 033,8**	**-4 786,7**
ln L()	**-4 117,4**	**-1 957,5**	**-3 317,8**	**-3 186,2**
Chi-Sq.(14)	**4 191,0**	**2 497,0**	**3 434,9**	**3 201,0**

Table 10: Regression results for annual reports from 1993-1996
z- values are reported in parenthesis
*) Indicates dummy variable
XX) Indicates that variable was dropped from regression by STATA

5.4 Interpretation of results

Table 10 presents the results. As can be seen, the variable coefficients have statistical significance, and they are relatively stable. t=1994 appears to be an exception. The reason is probably the general lack of annual reports for companies of 0-1 year. These estimates should therefore be given less weight in the analysis.

A general observation is that most of the coefficients also have the expected signs. Increasing solidity and revenues lead to decreased bankruptcy risk. Increased interest burden increase the risk of bankruptcy; young companies go more easily bankrupt, and lacking annual reports increase bankruptcy risk. In addition, one can see that restaurants are more risky than average and real estate is less risky than average.

Comparing only the regression for t=1993,1995 and 1996, the coefficients seem stable. The confidence intervals at 95% level are overlapping, indicating that the system is stable over time[63]. For the last regression, t=1996, interest burden seem less important in explaining bankruptcies year t=1998. One factor explaining this might be the low level of interest rates in Norway for the period 1996-Q3 1998. It might be that there is some non-linearity that make interest burden unimportant when the ratio falls below some level.

Concerning the dummy variables used for adjustment, at first, these could be interpreted as having the wrong sign. The initial interpretation could be that variables with an initially too high revenue reserve score should have decreased bankruptcy risk. However, these dummy variables need not have a clear-cut interpretation. An alternative interpretation of expected sign could be as follows: If a majority of the truncations have taken place due to a smaller than average denominator, rather then a “greater than average” nominator. Then the variable is actually more risky than what appears from having a high revenue reserves ratio, hence the dummy coefficient has the correct sign. Most probably are both of the explanations correct, this give not clear-cut interpretation of the signs of the dummy variable coefficients. This indicate that financial ratios have some weaknesses when used for sorting purposes.

[63] The time series is very short to make this fully credible, but the tendency is present.

Risk prediction

One way of testing the properties of the estimated models is to check how predicted risks are distributed. Each year can be divided into two sub-samples, the ones that go bankrupt and the ones that survive. This is done in Table 11 to Table 14. Each table lists the predictive ability for a regression model from year *t*, applied on balance sheets from a different year. Investigating predicted risks for the bankrupt and non-bankrupt separately. It can be seen that that the bankrupt group has a 10 times higher mean of predicted risks compared with the non-bankrupt. The bankruptcy group also has a much larger standard deviation making this group less homogenous. In the appendix, there is a graphical presentation of how the different estimation models correlate, indicating a strong stability between the different estimation models.

Mean predicted risks of bankruptcy in 1995 using annual reports from 1993				
	Non-bankrupt	**# obs.**	**Bankrupt**	**# obs.**
Prediction using 1993 estimates	0,014	77 696	0,159	1 201
	(0,043)		(0,116)	
Prediction using 1994 estimates	0,008	77 627	0,104	1 200
	(0,030)		(0,084)	
Prediction using 1995 estimates	0,013	77 642	0,153	1 200
	(0,043)		(0,113)	
Prediction using 1996 estimates	0,010	77 627	0,118	1 200
	(0,033)		(0,089)	

Table 11: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1993 data to predict bankruptcy in 1995.
Standard deviation in parenthesis

Mean predicted risks of bankruptcy in 1996 using annual reports from 1994				
	Non-bankrupt	**# obs.**	**Bankrupt**	**# obs.**
Prediction using 1993 estimates	0,014	72 638	0,184	544
	(0,043)		(0,104)	
Prediction using 1994 estimates	0,007	72 613	0,121	544
	(0,028)		(0,080)	
Prediction using 1995 estimates	0,012	72 629	0,171	544
	(0,040)		(0,099)	
Prediction using 1996 estimates	0,010	72 613	(0,14)	544
	(0,032)		(0,080)	

Table 12: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1994 data to predict bankruptcy in 1996.
Standard deviation in parenthesis

Mean predicted risks of bankruptcy in 1997 using annual reports from 1995				
	Non-bankrupt	**# obs.**	**Bankrupt**	**# obs.**
Prediction using 1993 estimates	0,011 (0,037)	81 297	0,149 (0,116)	916
Prediction using 1994 estimates	0,005 (0,024)	81 294	0,097 (0,083)	916
Prediction using 1995 estimates	0,010 (0,035)	81 312	0,144 (0,112)	917
Prediction using 1996 estimates	0,008 (0,028)	81 294	0,111 (0,089)	916

Table 13: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1995 data to predict bankruptcy in 1997.
Standard deviation in parenthesis

Mean predicted risks of bankruptcy in 1998 using annual reports from 1996				
	Non-bankrupt	**# obs.**	**Bankrupt**	**# obs.**
Prediction using 1993 estimates	0,013 (0,042)	87 701	0,159 (0,113)	849
Prediction using 1994 estimates	0,007 (0,028)	87 689	0,104 (0,083)	849
Prediction using 1995 estimates	0,012 (0,040)	87 706	0,152 (0,110)	849
Prediction using 1996 estimates	0,009 (0,031)	87 689	0,119 (0,087)	849

Table 14: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1996 data to predict bankruptcy in 1998.
Standard deviation in parenthesis

Financial distress literature commonly apply a “P>0,5 ⇒classified bankrupt” sorting rule. As can be seen from the table above, this rule would not prove very useful as most observations are far below this measure. A more appropriate, but ad hoc, value can for example be set at two standard deviations above the predicted risk for companies not going bankrupt, this measure ensures 95% correct classification of non-bankrupt companies.

Another way of evaluating the results is to examine its aggregation properties. The idea is to check if the model gives a correct description of the business cycle. With access to an almost complete population, the model should be able to describe the aggregated economy.

The sum of predicted risks for year *t* gives an indication of the aggregated level of bankruptcy year *t+2*[64]. By using the different estimations to make predictions about the full data set, it is possible to cross-tabulate the performance of the individual models/estimations.

Estimation year	Predicted bankruptcies in year 1995	1996	1997	1998
Actual*	1201	544	917	849
Official statistics**	2195	2141	2054	2493***
Estimates based on 1993 reports	1198	1032	990	1198
Estimates based on 1994 reports	653	542	512	634
Estimates based on 1995 reports	1130	945	915	1108
Estimates based on 1996 reports	856	730	701	847

Table 15: Prediction of aggregated bankruptcies using the regression estimates based on the 1993-1996 annual reports
*) Taken from the data set
**) Source: Bank og kredittstatistikk, various editions
***) Source: Dun and Bradstreet,
Note: The reason that the within year predicted numbers are higher than the actual number in the data set is that observations bankrupt in year *t+1* are not taken out of the prediction sample, compared with the estimation sample.

Table 15 reports the aggregation properties of the model estimates. Comparing the estimates with the numbers available in the data set do all the models over-predict bankruptcies in 1996. This prediction is based on the 1994 annual reports, which lack a majority of young companies. In addition, the majority of bankruptcies normally happen among firms younger than 5 years[65]. When lack of a large portion of this group does not have a significant impact on the estimates, the rational explanation is that age has less weight in the estimated model compared with what it should have. I interpret the over-prediction of 1996 bankruptcies as an under-valuation of age.

To find possible explanations to the effects of age in the system, age needs deeper investigation. The empirical results give an indication to presence of a minimum point of the age-bankruptcy relationship. Using the means of the age coefficients together with the results in (5.4), risk is minimised with respect to age when a company is approximately 65 years old:

---

[64] The sum of predicted probabilities is simply the expected value of the data set due to the specification of the model: $\sum_{i=1}^{N}\left[p_i \cdot 1 + (1-p_i)\cdot 0\right]$ where N is the total population.

[65] See table for distribution of bankruptcies, subject to age, in the appendix for details.

$$\frac{\partial P}{\partial (age)} = 0 \Leftrightarrow age = -\frac{\hat{a}_1}{2\hat{a}_2}$$
$$age = -\frac{-0{,}0235 * 100}{0{,}018 * 2} \cong 65\text{years} \tag{5.6}$$

In the data set, there are approximately 1700 observations each year that are older than 65 years. This group has only 3-8 bankruptcies each year. Low number of bankruptcies among old companies indicates a low risk of bankruptcy among older firms, and higher risk among young firms. In turn, this gives support to the selection theory of Jovanovic (1982). The capital vintage theory predicts a higher rate of exit for older firms, this is not supported by the results. This lack of support for the vintage capital theory probably comes from the lack of information on the age of capital equipment. I.e. age is decreasingly important in determining the risk of bankruptcy. In comparison to Salvanes and Tveterås (1999), they find the equivalent point at 15-18 years. Klette and Mathiassen (1996) find the equivalent point at 12-14 years[66].

There are two indications that the estimated age effect is too small. The first is that none of the estimated models seems to be able to evaluate the outcome of the 1994 observation correctly. The 1994 observations have half of the normal rate of registered bankruptcies and this year have a lack of young companies, i.e. companies aged 0 and 1 year. The second is the comparison of to previous empirical results. Compared with earlier studies, this study find that age is relatively less important in determining the risk of bankruptcy. The probable explanation is the sample selection bias. Decomposition of Statistics Norway's figures for bankruptcies by age groups[67] shows that approximately 25% of all company bankruptcies are among companies younger than 2 years and therefore unavailable in the data set due to the 2

---

[66] Salvanes and Tveterås (1999) and Klette and Mathiassen (1996) investigate the life span of the individual plant, with exit as the dependent variable. As exit includes mergers and voluntary liquidations, it is a more general term compared to bankruptcy and the results are therefore not directly comparable. Still, it is reasonable to believe that the results should indicate the same strength of the age effect. The different phenomena are highly related: Liquidation will happen if the company has no prospects for future profits and it has good routines for monitoring. An acquisition will often happen if a company can be bought at a discount. Bankruptcy will then happen with companies that no other company has interest in or the ones that have the weakest management routines.

[67] See table for distribution of bankruptcies, subject to age, in the appendix for details.

year prediction horizon. When the tendency is that 15% of the data set has 25% of the bankruptcies, there is reason to believe in incompleteness in the sample[68], i.e. that the data does not portray the youngest firms properly.

5.5 Out of sample tests

While the regression results seem quite stable, all of the years come from the same part of a business cycle. As the estimation period was a reasonably stable period of growth in Norway, the expectation is that the results are similar. Presence of the shakeout effect described by Caballero and Hammour (1994,1996) makes estimations from economic growth periods less useful to analyse periods of economic decline/recession.

To see how the results hold for years that are at different stages of the business cycle, the model was tested on data for 1989-1991. The years 1991 - 1993 saw the highest level of bankruptcies in Norwegian history. The real test of the estimated model is to see how it performs in this period of economic downturn. The observations from the period 1988-1991 cannot be directly compared with the estimation years due to the tax reform of 1992 discussed earlier. Some of the entries in every balance sheet were recalculated to incorporate the major effects of the reform.

To take the effects of the 1992 tax reform the data for 1989-1991 was modified. The idea is to recalculate the balance sheets as if the tax reform had already been implemented. This method has some faults, but 1989-1991 is the only period of recession where data are available. Recalculations were done as follows: the entry for conditional equity capital was split at a 60-40 rate on revenue reserves and long-term debt. This is a very rough measure, but it makes the mean level of revenue reserves approximately equivalent to the level of the data after 1992.

The tax reform also changed the way taxes were calculated. Due to the change in the marginal tax rate, the government allowed companies to shift income from 1991 to 1992 in an attempt to split the costs of the reform. While important for the individual company, these issues were not incorporated in the analysis.

[68] There exists estimation methods that adjust for sample selection, see for example Heckman (1979). These methods depend on information from the population that has been left out. In this study, there are no available information on the youngest companies, rendering the proposed methods useless.

Test results

The estimated models from 1993-1996 were used to predict bankruptcy risk for individual companies in the years 1989-1991. The models produce slightly differing results.

As was done for the estimation samples, it is possible to examine the distribution of predicted risks on bankrupt and surviving companies. All of the models generally assign a much higher bankruptcy risk to the companies that actually went bankrupt, compared with the ones that survived. Another interesting thing to note is that the mean is increasing for these years, which also is the trend for the bankruptcy ratio. These results are summed up in Table 16 to Table 18 below:

Feil! Ugyldig kobling.

Table 16: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1989 data to predict bankruptcy in 1991.
Standard deviation in parenthesis

Feil! Ugyldig kobling.

Table 17: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1990 data to predict bankruptcy in 1992.
Standard deviation in parenthesis

Feil! Ugyldig kobling.

Table 18: Mean predicted risks for companies observed as non-bankrupt or bankrupt, using 1991 data to predict bankruptcy in 1993.
Standard deviation in parenthesis

Table 19 shows what the estimations predict on the aggregated economy. If the predictions are compared with the actual numbers that can be found in the SEBRA data set, one finds that all of the models underpredict by a substantial margin. For the recession period, the empirical models developed are unable to find the correct level of bankruptcy risk, while they retain the ability to sort companies correctly. This is in accordance with shakeout effects, i.e. there will be a higher-than-normal rate of company exit during recessions as suggested in Aghion and Howitt (1992) and Caballero and Hammour (1994,1996).

Predicted bankruptcies in year			
Estimation year	1991	1992	1993
Data set appearances*	883	1881	1768
Official statistics**	3769	4446	3859
Estimates based on 1993 reports	585	1040	1061
Estimates based on 1994 reports	313	570	574
Estimates based on 1995 reports	522	954	971
Estimates based on 1996 reports	404	732	747

Table 19: Cross-tabulation of estimated aggregated bankruptcy predictions
*) The number of bankruptcies present in the data set.
**) Source: Statistics Norway; these numbers include unlimited liability companies, as the groups were not separately accounted for before 1994

The 1993-1997 sample supports the possibility of a U-shaped relationship between age and bankruptcies. Investigating 1989-1992 sample show that the relationship is not significant in explaining bankruptcies in 1991 and 1992, but it is significant in explaining bankruptcies in 1993[69]. This can be kept in relation to the shakeout effects described by Aghion and Howitt (1992) and Caballero and Hammour (1994,1996). They predict that companies will exit as a result of low productivity. It is reasonable to believe that age is stronger related to the learning effect described by Jovanovic (1982), than a measure for productivity. When age lose predictive power during the recession, this is an indication of another effect is present during recessions compared with booms. The low significance of age during the recession period can therefore be interpreted as dominance of the shakeout effect during recessions. No other explanations have been investigated to answer why the U-shaped effect of age on bankruptcies is disrupted when the economy moves into a recession[70]. Salvanes and Tveterås (1999) are unable to find evidence of both the U-shaped effect and the shakeout effect simultaneously.

Classification matrixes

In earlier literature it is common to set up matrices to identify how well a proposed model performs. It is most commonly used in multivariate discriminant analysis studies where it becomes relevant to what extent the model predicts incorrectly, i.e. classification of bankrupt companies as solvent or solvent companies as bankrupt[71]. In the probit framework, a cut-off

---

[69] The estimation results for the period 1989-1991 is reported in the appendix.

[70] One example of this could be that bankers change mentality to risk during a recession.

[71] Also known as error of type I and type II

point for predicted risk at 50% is the commonly used to investigate the classification properties. This is not necessarily a meaningful cut-off value as the number of bankruptcies is very low in this investigation. This gives dominance to the constant element in the level of risk and results in all observations being predicted at relatively low risk. To compare the performance of this model with other studies I have therefore made an ad hoc choice for comparison. Using approximately two standard deviations from the mean of predicted risks for non-bankrupt companies give a cut-off value of $Pr_{1993\ model}$ *(bankruptcy year t+2)=10%*. Boyes *et al.* (1989) state that the use of cut-off values generates a too narrow measure for credit risk assessment[72]. Table 20 gives a summary of the results. Tables showing the actual numbers of classification can be found in the appendix.

Table 20 shows that the results presented in this thesis are at par with earlier studies. The apparent weakness in correctly predicting bankrupt companies, but this is dependent on the decided cut-off value. Reducing the cut-off value will increase correctly predicted bankruptcies while reducing the number of correctly predicted solvent companies.

Rates of correct prediction for this and previous studies				
Study	**Corr. prediction of survivors**	**Corr. prediction of bankrupt**	**Total**	**# Obs**
This, t=1989	96,8	55,5	96,1	46 204
This, t=1990	96,1	65,0	95,3	67 034
This, t=1991	96,2	72,3	95,6	71 800
This, t=1993	95,5	68,6	95,1	78 897
This, t=1994	95,2	68,6	95,1	73 182
This, t=1995	96,7	65,4	96,4	82 213
This, t=1996	95,6	70,0	95,4	88 550
Altman *et al.* (1977)*	93,1	84,9	89,0	111
Altman *et al.* (1994)*	93,6	89,1	91,4	302
Bardos (1998)**	77,3	78,1	--	38 734
Bardos (1998)***	73,4	70,3	--	33 879
Ohlson(1980)	--	--	96,1	2 163
Olsen(1991)*	86,7	73,3	80,0	60
Platt & Platt(1990)*	88,0	91,0	90,0	68
Platt & Platt(1991)*	82,0	91,0	86,0	182

Table 20: Comparison of the results of this thesis and previous studies.
*) Results for a balance sample
**) Prediction on 1995 data

72 The argument is that credit lenders are interested in profit maximisation rather than default classification. A simple classification matrix does therefore not contain enough information for the bank. A continuous description of risk is therefore gives a more true to reality.

***) Prediction on 1991 data
-- Indicates that the number is not reported in the study
Note: For this study, the cut-off value is "*predicted risk>10%*". Cut-off values are generally not directly comparable between studies as they depend on the model specification/assumptions.

The most natural study to compare the results of this thesis with is Bardos (1998). This is the model employed in the French Central Bank. This thesis has apparently a higher classification power when compared to Bardos (1998), and it has a lower ability to classify correctly compared with studies using balance samples. The differences between usage of balanced[73] and unbalanced samples are significant. The explanation lies in differences in sample selection; unbalanced samples and balanced samples are probably not directly comparable. Zmijevski (1985) suggests that the process of selecting observations for a balanced sample makes the data predisposed to pattern recognition of bankrupt/non-bankrupt.

Change of the cut-off point can increase the number of correctly classified bankrupt cases, but will decrease the number of correctly predicted solvent companies. The preference for a high degree of correctly classified solvent companies comes from the fact that this group is much larger than the bankrupt group. A low percentage of correctly predicted solvent companies would therefore have a larger impact on macro analysis relative to a low number of correctly predicted bankrupt companies.

[73] A balanced sample has an equal proportion of distressed companies and non-distressed companies.

6 SUMMARY OF RESULTS AND CONCLUSION

In this thesis, an empirical model for estimating the risk of bankruptcy is presented. Using microeconomic information, I combine ideas from financial analysis and industrial organisation theory and find a set of stable variable coefficients that are stable over time. The result is a model that is useful in describing bankruptcy risk in the Norwegian business sector. The estimates show that the risk of bankruptcy increases when interest burden as a share of cash-surplus increase. Furthermore, the risk of bankruptcy decreases when either revenue reserves as a share of debt or income as a share of debt increases. Another strong indicator for bankruptcy is the lack of an annual report for the year before bankruptcy occurs.

The results support the idea that companies are learning entities, proposed in the selection model of Jovanociv (1982). Support for this learning process is supported with age being relatively much more important in determining bankruptcy for young companies, compared with older companies. The relationship between age and bankruptcy risk is U-shaped, i.e. the impact of age on risk of bankruptcy is not linear. The estimated model may under-emphasise the age effect due to under-representation of young companies that quickly go bankrupt. This under-representation of young firms that go bankrupt make the impact of company age appear less important in predicting bankruptcy then what it actually might be. The data does not contain information to either support or reject the vintage capital of Johansen (1959) or effects of financial structure on bankruptcy proposed on incomplete contracts theory (See Hart (1995)).

It was not found any good indication for significant differences in risk between the different segments of the economy. This is with the exception of companies operating in the real estate business, with a lower risk of bankruptcy, and companies in the restaurant business that have a higher risk of bankruptcy. Furthermore, the results do not support the existence significant differences in financial ratios for different sectors/segments of the economy as suggested by Platt and Platt (1990,1991). The findings indicate that all companies seem to be affected by financial ratios equivalently regardless of what the company produces.

The results from specification testing can be interpreted in accordance with the shakeout theory of Aghion and Howitt (1992) and Caballero and Hammour (1994, 1996). This is based on the observation the model retains its sorting properties on a holdout sample, but it is unable to describe the level of risk for the holdout sample properly. All predictions are too low for the macroeconomic downturn in 1989-1991, i.e. the number of bankruptcies in the economy surges when the economy enters a recession. In the holdout sample, the estimations show that age is unable to have any power to explain bankruptcies. The shakeout theory suggests that productivity becomes more important in determining exit during downturns compared with upturns. When age lose the ability to predict bankruptcy during the downturn this is interpreted as an indication that an unobserved measure for productivity becomes the important determinant.

Ratios created from the annual balance sheet retain power as determinants of bankruptcies. This stability in the coefficients give reason to believe that financial ratios are useful in predicting bankruptcies, but there is both room and the need for the inclusion of macroeconomic indicator to improve results in future bankruptcy prediction models.

BIBLIOGRAPHY

Aghion, P. and P. Howitt (1992): «A model of growth through creative destruction», *Econometrica*, vol. 60, nr 2, p.323-351

Aldrich, J.H., and F.D. Nelson (1984): *Linear Probability, logit and Probit Models*, Sage University Paper 45, Sage Publications,

Altman, E.I. (1968): «Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy», *Journal of Finance*, September, p.589-609

Altman, E.I., H. Frydman and D. Kao (1988): "Introducing recursive partitioning for financial classification: The case of financial distress» *Studies in Banking and Finance* 7, p. 197-223

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

Altman, E.I. and P. Naranayan (1997): "An international survey of business failure classification models», *Financial Markets, Institutions and Instruments*, Vol.6, No.2. p. 1-57

Altman, E.I. and A. Saunders (1998):"Credit risk measurement: Developments over the last 20 years", *Journal of Bankring and Finance*, Vol 21, p. 1721-1742

Andersen A. and T. Halvorsen (1992): «Konkursprediksjon på grunnlag av regnskapsdata (Using accountacy data for bankruptcy prediction)», Graduation thesis, Graduate school of Business, Tromsø

Argenti, J (1976): *Corporate collapse, the causes and symptoms,* McGraw-Hill, London,

Bardos M. (1998): «Detecting the risk of company failure at the Banque de France», *Journal of Banking & Finance*, Vol 22, p 1405-1419

Bardos, M. and W. Zhu (1997): "Comparison of discriminant analysis and neural networks: Application to the detection of failure", in *Bio-metric approaches in management science, Kluwer Academinc publishers*, p. 1-25

Basle Committee on Banking Supervision (1998): "Progress report, fact finding on credit risk modelling", Bank of International Settlements 98/62,

Beaver, W. (1966): «Financial ratios as predictors of failure», J*ournal of Accounting Research*, p.77-111:

Ben-Akiva, M., S.R.Lerman (1985): *Discrete choice analysis- Theory and Application to Travle Demand*, MIT Press, Cambridge, USA,

Boyes, W.J., D.L. Hoffman, S.A. Low (1989): "An econometric analysis of the bank credit scoring problem", *Journal of Econometrics*, 40, p.3-14

Bibliography

Blanchard, O. and P. Diamond (1990): «The cyclical behaviour of the gross flows of U.S. workers», *Brookings Papers on Economics Activity*, nr. 2. p.85-155

Blevins, C., H. Lehment, E. Sjøvoll (1997): «Gesamtwirtschaftliche Bestimmungsgründe der Insolvenzentwicklung in Deutschland (Macroeconomic determinants in the development of German bankruptcies)», *Kiel Working Papers*, Working Paper No.842, Kiel,

Breiman L., J.H. Friedman, R.A. Olshen and C.J. Stone (1984): *Classification and regression trees* Wadsworth, Belmont, USA,

Bresnahan, T.F. and D. M.G. Raff (1991): «Intra-industry heterogeneity and the Great depression: The American motor vehicles industry 1929-1935», *Journal of Economic History*, vol. 51, nr. 2, p.317-331

Bresnahan, T.F. and D. M.G. Raff (1992):»Technological Heterogeneity, Adjustment costsand the dynamics of plant shut-down behaviour: The American motor vehicle industry in the time of the Great depression» Mimeo, Stanford University

Bruflot, G. (1993): «Konkursprediksjon - en litteraturstudie (Bankruptcyprediction- a litterature survey)», SNF Workingpaper no. 14/93, SNF, Norway,

Caballero R.J. and M.L. Hammour (1994): «The cleansing effect of creative destruction», *American Economic Review*, Vol 84, no. 5, p.1350-1368

Caballero R.J. and M.L. Hammour (1996): «On the timing and efficiency of creative destruction», *Quarterly Journal of Economics*, August, p. 805-852

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

Eisenbeis, R. (1977): «Pitfalls in the application of discriminant analysis in business and economics», *The Journal of Finance*, June, p. 875-900

Eklund T. (1988): «*Konkursindikator - et nyttig analyseverktøy* (Are bankruptcy indicators useful for analysis?)" Graduation thesis (HAS), NHH.

Feller W. (1968): «*An Introduction to Probability Theory and its Applications*, Wiley, New York,

Fisher, R.A. (1936): « The Use of Mulitiple Measurement in Taxonomic Prolems» *Annals of Eugenics* 7, p.179-188

Fjeld, S., E. Gaaseide and J. Stensrud (1994)*: Regnskapsstatistikk 1991-1992: Private ikke-finansielle foretak med begranset ansvar*(Accounting statistics 1991-1992)» Notater, Statistics Norway,

Gessner G., W.A. Kamakura, N. K. Malhorta and M.E. Zmijewski (1988): «Estimating models with binary dependent variables: Some theoretical and empirical observations», *Journal of Business Research*, Vol.16, no. 1, p. 49-65

Ghemawat, P. and B. Nalebuff (1985): «Exit», *Rand Journal of Economics*, 16, p.184-194

Gjesdal, F. (1995): «Short-term prediction of bankruptcy based non-financial indicators» SNF-report 48/95, SNF,

Greene, W. (1993): *Econometric Analysis*. Macmillan Publishing Company, New York,

Grossman, G. and E. Helpman (1991): *Innovation and growth in the global economy,* Cambridge, MA, MIT press,

Hart, O. (1995): *Firms, contracts, and financial structure*, Oxford University Press, Oxford

Heckman, J.(1979):"Sample selection bias as a specification error", *Econometrica*, Vol 47, pp. 153-161

Howieson B. (1991): «A Security analyst's action recommendations: An application of recursive partitioning to modelling judgement» *Australian Journal of Management,* Vol. 16, 2, december, p. 165-185

Jensen, M.C. (1993):"The modern industrial revolution, Exit, and the failure of internal Control systems", *Journal of Finance*, Vol. 48, P.831-880

Johansen, L. (1959): "Substitution versus fixed production coefficients in the theory of economic growth: a synthesis" *Econometrica*, vol 27, p.157-176

Jones, F.L. (1987): «Current techniques in bankruptcy prediction», *Journal of Accounting Litterature*, p. 131-164

Jovanovic B. (1982): «Selection and the evolution of industry», *Econometrica*, Vol. 50, no.3, p. 649-670

Karnov (1996): *Norsk kommentert lovsamling* (Comments to the Norwegian code of Laws), Karnov, Oslo, p. 2195-2241

Keasey K, and R. Watson (1986): «The State of art of small firm failure prediction: Achievements and prognosis», *International Small Business Journal*, nr. 4, p. 11-29

Klette, T.J. and A. Mathiassen (1996): *Vekst og fall blandt norske industribedrifter: om nyetablering, nedlegging og omstiling* (Turnover and growth among Norwegian manufacturing plants), Social and Economic Studies nr 95, Statistics Norway, Oslo,

Knivsflå, K.H. (1997): "Tek det lang tid før verksemder i økonomisk krise vert slegne konkurs?}» SNF working paper nr 43/1997, SNF,

Langli, J.C. (1994): «Konkurskriminalitet: En empirisk analyse av Aksjeselskaper som har gått konkurs» *Forskning om økonomisk kriminalitet*, nr 17 Norges Forskningsråd, Handelshøyskolen BI, sandvika,

Lo, A.W. (1986): «Logit versus discriminant analysis: A specification test and application to corporate bankruptcy», *Journal of Econometrics*, Vol. 31, p. 151-178

Martin, S. (1993): *Advanced Industrial Economics*, Blackwell Publishers, Cambridge, MA

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

Olsen J.P. (1991): “En empirisk analyse av modeller for konkursprediksjon (An empirical study of models predicting bankruptcy”, Graduation thesis (HAS), NHH,

Platt H.D. and M.B. Platt (1990): «Development of a class of stable predictive variables : The case of bankruptcy prediction», *Journal of Business Finance & Accounting*, Vol. 17, no. 1, p.31-51

Platt H.D. and M.B. Platt (1991): “A note on the use of industry-relative ratios in bankruptcy prediction” *Journal of Banking & Finance*, Vol. 15, p.1183-1194

Pratt, J.W. (1981):”Concavity of the Log-likelihood”, *Journal of the American Statistical Association*, , Vol. 76, p. 137- 159

Maddala, G.S. (1983): *Limited-Dependent and qualitative variables in Econometrics*, Cambridge University Press, Cambridge, USA

NOU (1993): *Etterkontroll av konkurslovgivningen m.v.* (Inspection of the Norwegian bankruptcy legislation), NOU, OSLO,

Salvanes, K. and R. Tveterås (1998): “Firm exit, Vintage effect and the Business cycle in Norway» Discussion paper nr 2/99, NHH, Norway

Scott, J. (1981): «The probability of bankruptcy: A comparison of empirical predictions and theoretical models», *Journal of Banking and Finance*, p. 317-344

SND (1995): “System for monitoring loans to the business sector», SND rapport 5-1995, Oslo, Norway

STATA Corporation (1997): *STATA reference manual, release 5*, Vol. 3, Stata Press, College Station, Texas, USA

Tirole, J. (1989): *The Theory of Industrial Organization*, The MIT press, Cambridge, MA, USA

Wilcox J.W. (1971): « A simple theory of financial ratio as predictors of failure» *Journal of Accounting Research*, autumn, p.389-395

Wilcox J.W. (1976): “The Gambler's ruin approach to business risk», *Sloan Management Review*, Fall, p.33-46

Zavgren, C.V. (1983): «The prediction of corporate failure: The state of the art», *Journal of Accounting Literature*, nr. 2

Zmijewski, M.E. (1984): «Methodological issues related to estimation of financial distress prediction models», *Journal of Accounting Research*, Vol. 22, supplement, p.59-8

APPENDIX

Marginal effects of regressions:

Marginal effects of coefficients, bankruptcies year t+2				
Variable	**1993**	**1994**	**1995**	**1996**
Missing report year t+1,*	0,0881	0,0467	0,0719	0,0514
	(0,084)	(0,085)	(0.0666)	(0,083)
Revenue reserves>5*	-0,0034	-0,0004	-0,0020	-0,0018
	(0,801)	(0,822)	(0,834)	(0,872)
Revenue reserves>5*	0,0293	0,0005	0,0170	0,0214
	(0.0798	(0.447)	(0,081)	(0,091)
Revenue reserves<-1*	XX	XX	0,0065	XX
			(0,000)	
Self-financing	-0,0048	-0,0014	-0,0029	-0,0023
	(0,123)	(0.1490)	(0,140)	(0,146)
Self-financing>1*	0,0115	0,0011	0,0029	0,0022
	(0,062)	(0,060)	(0,063)	(0,066)
Self-financing<-1*	-0,0029	-0,0007	-0,0018	-0,0014
	(0,023)	(0,019)	(0,021)	(0,023)
Interest payments	0,0020	0,3100	0,0010	0,0005
	(0,035)	(0,309)	(0,308)	(0,302)
Interest payments>2*	-0,0032	-0,0008	-0,0019	-0,0015
	(0,070)	(0,066)	(0,066)	(0,065)
Interest payments<0*	-0,0026	-0,0006	-0,0014	-0,0010
	(0,022)	(0,020)	(0,023)	(.0238)
Restaurant*	0,0061	0,0008	0,0032	0,0012
	(0,023)	(0,021)	(0,023)	(0,024)
Real estate*	-0,0023	-0,0007	-0,0017	-0,0017
	(0,307)	(0,307)	(0,313)	(0,320)
Age	-0,0001	0,0000	-0,0002	-0,0001
	(10,656)	(11.97)	(10,860)	(10.7515)
Square of age	0,0001	0,0000	0,0002	0,0001
	(3,459)	(3.8134)	(3.364)	(3,235)
Age missing*	0,0036	XX	XX	XX
	(0,001)			
Observed P	**0,0153**	**0,0075**	**0,0112**	**0,0096**
Predicted P at x-bar	**0,0031**	**0,0006**	**0,0020**	**0,0015**

Figures in parenthesis are the mean of the relevant variable.
*) dF/dx is for discrete change of dummy variable from 0 to 1

Table 21: Marginal effects of the different variables

The z-values reported in chapter 5.

*) dF/dx is for discrete change of dummy variable from 0 to 1
Observed P is the sample portion of bankruptcy observations. Predicted P at x-bar indicates the predicted probability for bankruptcy for a hypothetical observation that has mean values.

Regression results for estimation on 1989-1991 data

Regression results for bankruptcies year t+2			
Variable	**1989**	**1990**	**1991**
Missing report year t+1	1,465	1,847	1,978
	(0,039)	(0,067)	(0,069)
Revenue reserves	-0,563	-0,404	-0,406
	(-0,011)	(-0,013)	(-0,013)
Revenue reserves>5*	2,224	1,361	1,254
	(0,006)	(0,006)	(0,006)
Revenue reserves<-1*	XX	XX	XX
Self-financing	-0,378	-0,286	-0,337
	(-0,004)	(-0,004)	(-0,005)
Self-financing>1*	0,170	0,294	0,196
	(0,001)	(0,002)	(0,001)
Self-financing<-1*	-0,939	-0,574	-0,547
	(-0,003)	(-0,005)	(-0,005)
Interest burden	0,293	0,208	0,109
	(0,006)	(0,006)	(0,003)
Interest burden>2*	-0,425	-0,471	-0,330
	(-0,004)	(-0,006)	(-0,003)
Interest burden<0*	-0,196	-0,596	-0,299
	(-0,002)	(-0,007)	(-0,003)
Restaurant*	0,041	0,236	0,230
	(0,000)	(0,004)	(0,004)
Real estate*	-0,410	-0,323	-0,362
	(-0,009)	(-0,010)	(-0,010)
Age	-0,002	-0,007	-0,018
	(-0,001)	(-0,003)	(-0,007)
Square of age	-0,004	0,001	0,015
	(-0,001)	(0,000)	(0,005)
Age missing*	4,262	14,570	36,815
	(0,001)	(0,003)	(0,007)
Constant	-2,404	-2,265	-2,313
	(-0,062)	(-0,090)	(-0,089)
pseudo R2	**0,298**	**0,371**	**0,430**
# obs.	**45 990**	**66 297**	**71 026**
ln L(0)	**-4 154,71**	**-8 537,15**	**-8 271,71**
ln L()	**2 918,62**	**-5 366,77**	**-471 595,56**
Chi-Sq.(14)	**2 472,17**	**6 340,75**	**7 111,52**

Table 22: Regression results for data from 1989 – 1991.
z- values are reported in parenthesis
*) Indicates dummy variable
XX) Indicates that variable was dropped from regression by STATA

Bankruptcies in the economy according to company age

Number of bankruptcies in the economy by age category					
Year	<2 years	2-5 years	>5 years	Unknown*	Total
1994	514	946	797	1377	3634
1995	559	914	750	1277	3500
1996	552	909	712	1285	3458
1997	538	848	709	1238	3333

* includes also personal bankruptcies

Source: Statistics Norway, Bank og Kredittstatistikk, Various issues

Table 23: Bankruptcies in the economy subject to the age of company

Report from the full model

Different model specifications were tested. Below are the regression results for a model including a dummy for observations that are bars, a liquidity measure and return on capital.

As a measure for size, the logarithm for revenues from operations was used. The variable measuring liquidity, *[Liquid assets - short-term debt]/ revenue from operations*. Another specification for liquidity tested was the current ratio. The current ratio[74] proved difficult to implement in a logical manner as many companies have registered either zero liquid assets or zero short-term debt. Different specifications either came up insignificant, or there were significant problems with causality. Return on capital is defined as *(revenue before extraordinary income and cost + interest burden)/ (Total debts + Equity).*

The LR test described in chapter 3 can be used for testing the effect of leaving a variable out of the maximum likelihood estimation. In this analysis, it was tested if the variables Dbar, Working capital, Working capital>50, Working capital<-15, Return on capital, Return on capital>1 and Return on capital<-1, and the logarithm of operations revenue are all equal to zero. This hypothesis of zero coefficients was rejected with 99.99% confidence.

The variable for size was left out from further analysis as the coefficient implies that bankruptcy risk increases is positively related to size of company. This result is not credible and is an indication of misspecification. A different hypothesis was then tested; H_2: the

[74] The current ratio is defined as liquid assets/short term debt

variables Dbar, working capital and return on capital are zero[75]. For the estimations on t=1994, 1995 and 1996 H_2 was rejected with respectively 33.5%, 81,9% and 81,6% confidence. The rationale for this test was that only the continuous variables are meaningful to use. The dummy variables are included only for adjustment purposes, rejecting the continuous variables make it rational to leave out also the respective dummy variables. This even though the LR test indicated that they are significantly different from zero.

[75] The dummy variables for both working capital and return on capital are kept in the regression.

Regression results for bankruptcies year t+2				
Variable	**1993**	**1994**	**1995**	**1996**
Missing report year t+1	1,598	1,725	1,553	1,493
	(49,83)	(35,59)	(46,04)	(43,38)
Revenue reserve	-0,460	-0,232	-0,393	-0,480
	(-13,79)	(-5,54)	(-10,33)	(-11,13)
Revenue reserves>5*	1,433	0,450	1,216	1,727
	(7,47)	(1,70)	(5,55)	6.894
Revenue reserve<-1*	XX	XX	0,725	XX
			(1,52)	
Self financing	-0,550	-0,589	-0,514	-0,493
	(-5,76)	(-4,41)	(-4,87)	(-4,56)
Self financing>1*	0,966	-0,563	0,586	0,538
	(6,30)	(2,26)	3.256	(2,97)
Self financing<-1*	-0,351	-0,608	-0,390	-0,351
	(-3,00)	(-3,43)	(-3,04)	-2.654
Interest burden	0,183	0,328	0,136	0,073
	(3,74)	(4,37)	(2,11)	(1,06)
Interest burden>2*	-0,219	-0,651	-0,283	-0,066
	(-2,00)	(-3,76)	(-1,48)	(-0,43)
Interest burden<0*	-0,325	-0,674	-0,237	-0,193
	(-2,92)	(-4,02)	(-1,74)	(-1,33)
Restaurant*	0,387	0,268	0,322	0,206
	(6,29)	(2,92)	(4,87)	(2,91)
Bar*	0,136	0,537	0,520	0,506
	(0,57)	(1,62)	(2,52)	(2,55)
Real estate*	-0,185	-0,272	-0,255	-0,313
	(-4,76)	(-4,71)	(-5,77)	(-6,72)
Age	-0,015	-0,020	-0,033	-0,024
	(-4,87)	(-4,55)	(-9,57)	(-7,15)
Square of age	0,01135	0,011	0,027	0,021
	(2,87)	(1,89)	(6,73)	(5,31)
Age missing*	0,332	XX	XX	XX
	(-0,80)			
Working capital	-0,001	-0,001	0,002	0,004
	(-0,39)	(-0,08)	(0,37)	(0,95)
Working Capital>50	-0,144	-0,144	-0,553	-0,395
	(-0,66)	(-0,04)	(-2,01)	(-1,49)
Working capital<-15	-0,030	-0,063	-0,065	-0,029
	(-0,27)	(-0,34)	(-0,51)	(-0,20)
Return on Capital	-0,219	-0,249	-0,091	0,124
	(-2,64)	(-2,07)	(-0,93)	(-1,25)
Return on capital>1	-0,335	-0,246	-0,445	-0,295
	(-2,34)	(-1,17)	(-2,69)	(-1,84)
Return on capital<-1	-0,651	-0,383	-0,414	-0,375
	(-7,49)	(-3,16)	(-4,44)	(-3,98)
Ln (income from operations)	0,072	0,063	0,035	0,062
	(8,47)	(5,26)	(3,75)	(6,47)
Constant	-2,973	-3,340	-2,621	-2,952
	(-40,49)	(-30,36)	(-33,01)	(-35,24)
pseudo R2	**0,354**	**0,401**	**0,350**	**0,346**
# obs.	**78 539**	**72 824**	**82 092**	**88 124**
ln L(0)	**-6 208,3**	**-3 205,7**	**-5 033,3**	**-4 786,2**
ln L()	**-4 010,7**	**-1 919,3**	**-3 271,3**	**-3 129,7**
Chi-Sq.(14)	**4 395,2**	**2 573,2**	**3 524,1**	**3 313,2**

Table 24: Regression results for explaining bankruptcies in year t+2 with balance sheets from year t. An extended model.

XX indicates that the variable was left out of the regression

*) Indicates dummy variable

Graphical presentation of model stability

As the size of data is large, a graphical presentation of the data is needed. Figure 2 to Figure 5 present the correlation between predicted risks using different year estimates. The graphs show how the estimations from different years correlate. Each graph portrays a separate year. The x-axis gives the predicted risks for observations for the year denoted over the graph, using the estimations from year 199x (denoted by p9x). The y-axis gives the predicted risks for the observations when using the estimations from year 199z (denoted by p9z). If there were 100% correlation between the two estimation years, each graph would be only a thin, 45^{o} line.

Some interesting details are found when examining the graphs. The first is that the separate estimations give highly correlated predictions, strengthening the notion of parameter stability between years as indicated earlier. The second is that the graphs including 1994 estimations have the worst fit/correlation. This is seen from the relatively large spread of the observations.

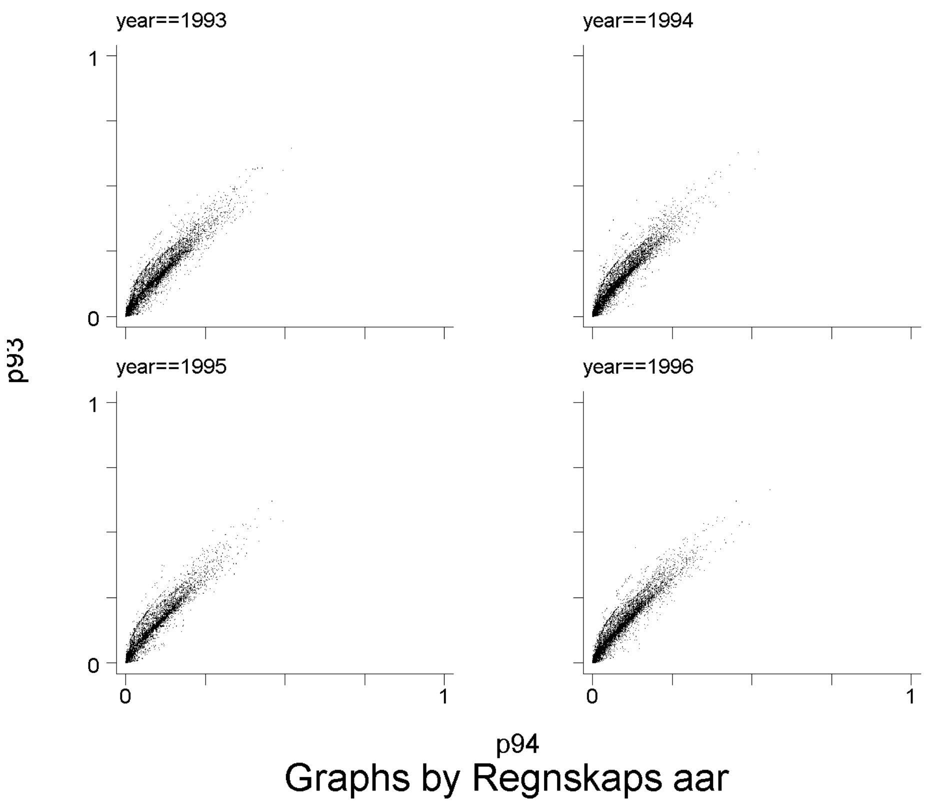

Figure 2: Plot of how predicted risks correlate using estimated models from 1993 (y-axis) and 1994 (x-axis). Each graph represents balance sheets from indicated year t to predict bankruptcy in year t+2.

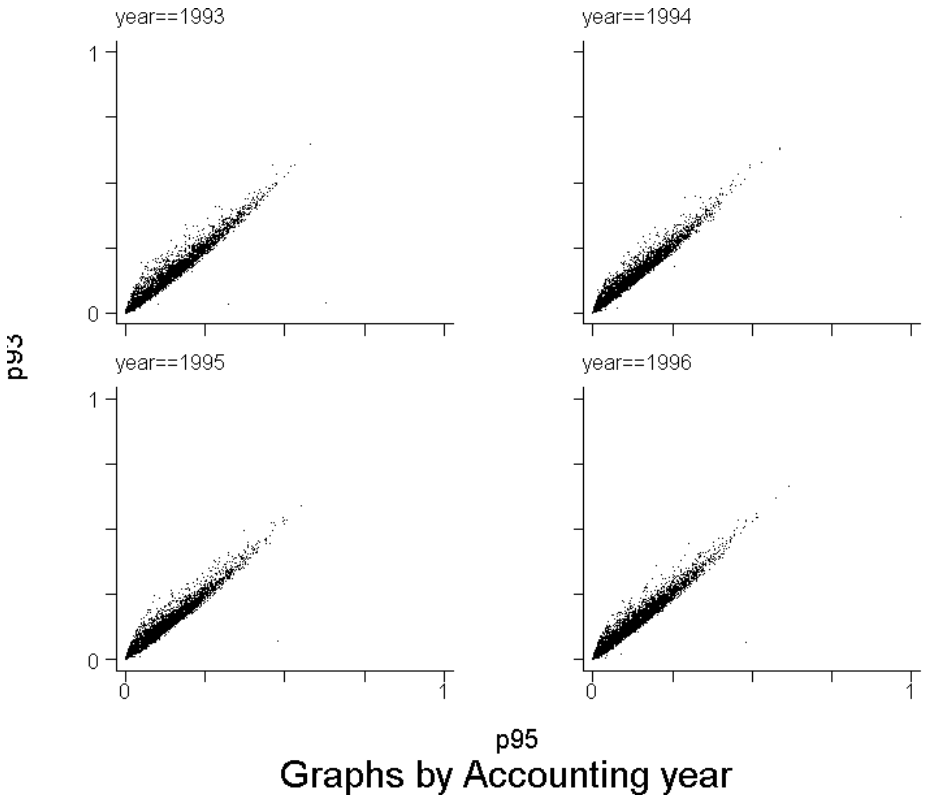

Figure 3: Plot of how predicted risks correlate using estimated models from 1993 (y-axis) and 1995 (x-axis). Each graph represents balance sheets from indicated year t to predict bankruptcy in year t+2.

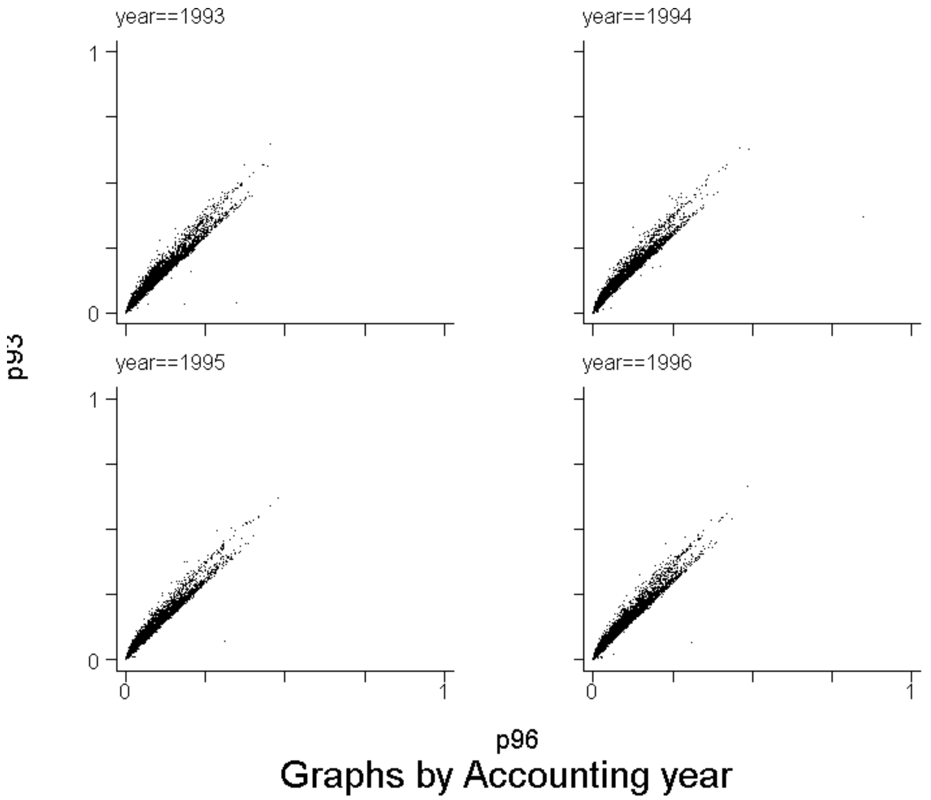

Figure 4: Plot of how predicted risks correlate using estimated models from 1993 (y-axis) and 1996 (x-axis). Each graph represents balance sheets from indicated year t to predict bankruptcy in year t+2.

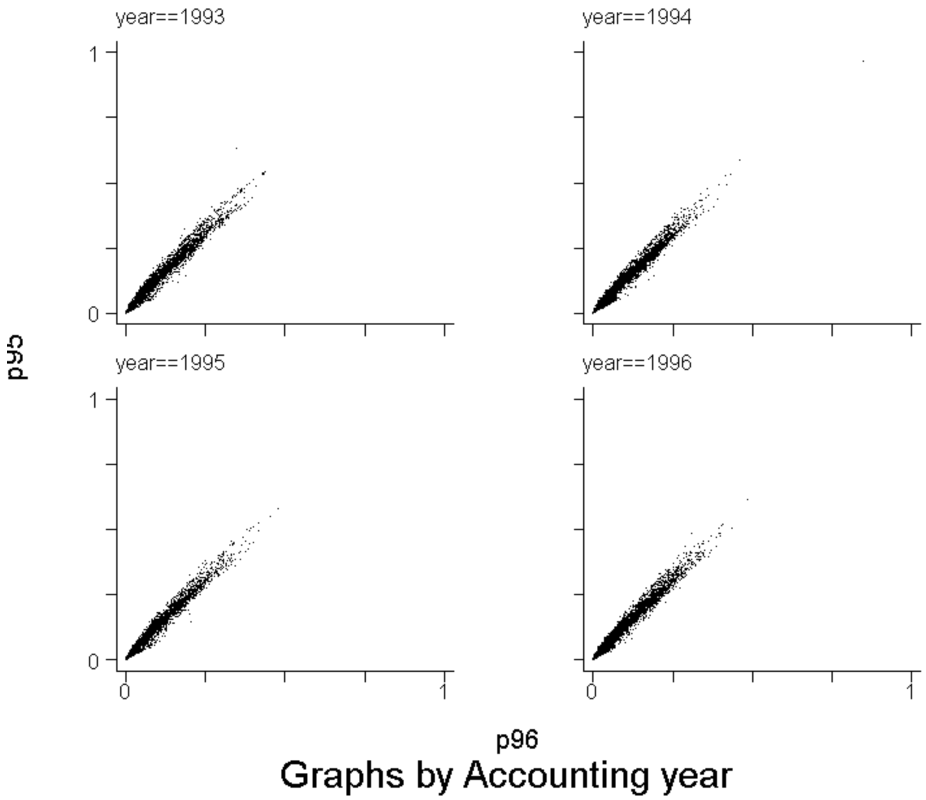

Figure 5: Plot of how predicted risks correlate using estimated models from 1995 (y-axis) and 1996 (x-axis). Each graph represents balance sheets from indicated year t to predict bankruptcy in year t+2.