DANMARKS NATIONALBANK

DANMARKS NATIONALBANK WORKING PAPERS 2004 • 17

Anne Dyrberg

Financial Market, Danmarks Nationalbank
and
Centre for Applied Microeconometrics (CAM), Institute of Economics, University of Copenhagen

Firms in Financial Distress: An Exploratory Analysis

18 June 2004

The Working Papers of Danmarks Nationalbank describe research and development, often still ongoing, as a contribution to the professional debate.

The viewpoints and conclusions stated are the responsibility of the individual contributors, and do not necessarily reflect the views of Danmarks Nationalbank.

As a general rule, Working Papers are not translated, but are available in the original language used by the contributor.

Danmarks Nationalbank's Working Papers are published in PDF format at www.nationalbanken.dk. A free electronic subscription is also available at this Web site.

The subscriber receives an e-mail notification whenever a new Working Paper is published.

Please direct any enquiries to
Danmarks Nationalbank, Information Desk, Havnegade 5, DK-1093 Copenhagen K Denmark
Tel.: +45 33 63 70 00 (direct) or +45 33 63 63 63
Fax : +45 33 63 71 03
E-mail:info@nationalbanken.dk

Nationalbankens Working Papers beskriver forsknings- og udviklingsarbejde, ofte af foreløbig karakter, med henblik på at bidrage til en faglig debat.

Synspunkter og konklusioner står for forfatternes regning og er derfor ikke nødvendigvis udtryk for Nationalbankens holdninger.

Working Papers vil som regel ikke blive oversat, men vil kun foreligge på det sprog, forfatterne har brugt.

Danmarks Nationalbanks Working Papers er tilgængelige på Internettet www.nationalbanken.dk i pdf-format. På webstedet er det muligt at oprette et gratis elektronisk abonnement, der leverer en e-mail notifikation ved enhver udgivelse af et Working Paper.

Henvendelser kan rettes til :
Danmarks Nationalbank, Informationssektionen, Havnegade 5, 1093 København K.
Telefon: 33 63 70 00 (direkte) eller 33 63 63 63
E-mail: info@nationalbanken.dk

ISSN (trykt/print) 1602-1185

ISSN (online) 1602-1193

Anne Dyrberg*

Firms in Financial Distress: An Exploratory Analysis

* I would like to thank Hans Christian Kongsted, Martin Browning, Mette Ejrnæs, David Lando, Jesper Rangvid and colleagues at the Danish Central Bank for commenting on this or earlier versions of the paper. I would also like to thank seminar participants at the DGPE-workshop, CAM seminar and at a seminar held at the Danish Central Bank.

RESUMÉ:

Danmarks Nationalbank introducerede sin årlige analyse af finansiel stabilitet i 2000. Formålet med analysen er at identificere de risici, som den finansielle sektor står overfor. Et væsentligt element i analysen af finansiel stabilitet er analysen af bankernes kreditrisiko, herunder af situationen i den ikke-finansielle sektor. Hovedformålet med denne artikel er at lave et redskab, der kan benyttes i analysen af den ikke-finansielle sektor eller, mere konkret, at opstille en model, der kan forudsige, hvilke virksomheder der ender i økonomiske vanskeligheder. Da virksomhederne i den ikke-finansielle sektor kan ophøre af forskellige årsager, herunder økonomiske vanskeligheder, frivillig lukning eller pga. opkøb mv., foreslås det at opstille en konkurrerende-risiko-model (”competing-risks model”).

Det økonomiske set-up bag en sådan model præsenteres og diskuteres, og modellen estimeres med udgangspunkt i danske data. Den empiriske analyse er baseret på en database af danske aktie- og anpartsselskaber, der eksisterede mellem 1995 og 2001. Databasen dækker omkring 30.000 virksomheder og mere end 150.000 observationer, og mere end 20 forklarende variable er inkluderet i analysen. Alle danske aktie- og anpartsselskaber analyseres, hvilket betyder, at der, givet tilskæring af database, estimeres på en repræsentativ sample.

Sammenlignet med den eksisterende litteratur er flere nye elementer introduceret i analysen. For det første skelnes mellem tre måder at ophøre på, nemlig virksomheder i økonomiske vanskeligheder, frivillig lukning og opkøb mv., og en konkurrerende-risiko-model estimeres. For det andet er der benyttet et stort antal proxy variable for uobserverbare faktorer såsom usikkerhed, evne og motivation. For det tredje giver den enestående database muligheden for at sammenligne forskellige specifikationer af kreditrisikomodeller. Den konkurrerende-risiko-model er sammenlignet med en pooled-logit-model (hvor alle ophørte virksomheder er modelleret i en gruppe) og en simpel konkursmodel (hvor de virksomheder der stopper på grund af økonomiske vanskeligheder er modelleret, og alle andre virksomheder er censorerede).

ABSTRACT:

Danmarks Nationalbank introduced its annual report on financial stability in Denmark in 2000. The purpose of the analyses is to identify risks currently faced by the financial sector. As the stability in the financial sector depends on the customers' financial circumstances, and as the majority of lending from Danish banks is granted to companies in Denmark, analyses of the development in the non-financial sector are crucial in a financial stability context.

The primary goal of this paper is to make a tool that can assist the regular analyses of the non-financial sector, namely to make a model that is able to predict the firms that end up in financial distress. As the firms in the non-financial sector may go out of business for various reasons (financial distress, voluntary liquidation, and because they are merged or acquired, etc.) the method of competing-risks models seems appropriate. To get point identification a parametric competing-risks model is suggested.

The parametric competing-risks model is estimated and the results are reported. The empirical analysis is based on a panel data set containing information on the whole population of Danish non-financial public limited liability companies (“aktieselskaber”) and private limited liability companies (“anpartsselskaber”) that existed between 1995 and 2001, covering around 30,000 firms and more than 150,000 firm-year observations. After application of certain criteria (e.g. exclusion of holding companies and financial institutions), the sample is representative. More than 20 explanatory variables are included in the estimations.

Compared to the existing literature this study introduces a number of novel elements to the empirical analysis.

Firstly, the empirical distinction between three modes of exit is developed, namely between firms in financial distress, voluntarily liquidated firms, and firms that merge with other firms or are acquired by other firms, and a competing-risks model is estimated. Secondly, a large number of proxies are used for inherently unobservable variables (e.g. uncertainty, ability and motivation). As is discussed in the paper, proxies are important. Thirdly, the extraordinary data set provides an opportunity to compare different specifications of credit risk models, and so the competing-risks specification is compared to a pooled logit model (where all exits are pooled) and to a simple financial distress model (where the exit to financial distress is modelled and all other firms are treated as censored).

1. Introduction

The aim of this paper is to present the set-up of an accounting-based credit scoring model and to estimate such a model using Danish data. The whole population of Danish public and private limited liability companies is included in the estimations. This means that after the application of certain criteria, e.g. exclusion of holding companies and financial institutions, the sample is representative. All sectors of the Danish economy are covered.

Models that can predict the firms that end up in financial distress are used by central banks and financial institutions. In central banks their predictions serve as input to financial stability assessments. Danmarks Nationalbank is one among many central banks, which have started publishing financial stability assessments on a regular basis, see e.g. Danmarks Nationalbank (2004). The purpose of the analysis is to identify the risks currently faced by the financial sector. The stability of the financial sector depends on the customers' health and so analysis with special emphasis on the banks and of how the financial sector is affected by the finances of companies and households are crucial. As the majority of lending from Danish banks is granted to companies in Denmark, it is primarily the developments in the Danish corporate sector that affects the banks. Accordingly a model that can predict the firms that end up in financial distress is of particular interest.

Individual financial institutions use credit-scoring models to assess the quality of a particular borrower. The topic of credit scoring has received renewed interest as Basel II opens up for the possibility that the credit institutions themselves can calculate the minimal capital requirements. According to Basel II the credit institutions have a choice of using a standard approach or one of two internal ratings-based approaches. Using either of the two internal ratings-based approaches, the credit institutions themselves must assess the probability that a borrower will default during the following year.[1] The model framework developed in this paper can be used in credit institutions to assess the probability that a borrower may default.

The literature on bankruptcy prediction is not new. The study of Beaver from 1966 is considered the pioneering work on bankruptcy-prediction models. The "theory" behind the model can best be explained within the framework of a "cash-flow". Beaver (1966:80) writes: "The firm is viewed as a reservoir of liquid assets, which is supplied by inflows and drained by outflows. The reservoir serves as a cushion or buffer against variations in the flows. The solvency of the firm can be defined in

---

[1] The Basel Committee is close to agreeing on the final content of the revised capital requirements, which should reflect more clearly the risks incurred by the individual credit institution. With the revised capital requirements the capital adequacy will to some extent be based on the credit institutions' own models, and the capital requirements will vary among credit institutions. For further details see Borup and Lykke (2003).

terms of the probability that the reservoir will be exhausted at which point the firm will be unable to pay its obligations as they mature (i.e., failure)”. Beaver (1966) uses a matched sample of 79 failed and non-failed firms. Using univariate discriminant analysis he shows that financial ratios can be used to predict corporate failure. Since the study of Beaver (1966), bankruptcy studies have been improved and refined. Altman (1968) uses a matched sample and introduces multivariate discriminant analysis. Ohlson (1980) suggests the use of the logit model, which among other advantages has the advantage that it is not needed to have a matched sample. Shumway (2001:101) criticises the approaches taken in the older studies, as they are single-period classification models: "By ignoring the fact that firms change through time, static models produce bankruptcy probabilities that are biased and inconsistent estimates of the probabilities that they approximate." Instead he proposes to use a hazard model.

Compared to the existing literature this study takes the analysis a step further. A number of novel elements to the empirical analysis are introduced.

Firstly, the empirical distinction between three modes of exit is developed, namely between firms in financial distress, voluntarily liquidated firms, and firms that have merged with other firms or are acquired by other firms, and so a competing-risks model is estimated. Most studies do not distinguish between exit types. The Altman (1968), Ohlson (1980) and Shumway (2001) studies focus on firms that go bankrupt. Other studies pool the various exit types, e.g. Bunn (2003) who uses the probit approach to model the probability that a company fails. One exception is Harhoff et al. (1998) who distinguishes between two (not three) exit modes. Harhoff et al. (1998) concludes that one should distinguish between firms that exit because of financial distress and firms that are voluntarily liquidated.

Secondly, a richer set of explanatory variables is included in the estimations compared to the explanatory variables that are usually included in estimations like this. Here age, the return on net assets, the solvency ratio, the short term debt to total assets and size are included as core variables, and on top of these variables a number of proxies (e.g. diversification and location dummies, and dummies which indicates the presence of critical comments from the auditors, whether the firm is an ultimate parent company, whether the company is owned by the public sector or a fund) and controls are included (e.g. sector affiliation dummies and year dummies). The proxies serve as proxies for variables that are inherently unobservable (uncertainty, ability and motivation). As is discussed in the paper, proxies are important.

Thirdly, the extraordinary data set provides an opportunity to compare different specifications of credit risk models, and so the competing-risks specification is compared to a pooled logit model (where all exits are pooled) and to a simple financial distress model (where the exit to financial distress is modelled treating all other firms as censored).

The conclusions are the following. First of all it is found that if the sign of the coefficients to the explanatory variables are of interest, then one should estimate the competing-risks model. This means that the conclusion in the Harhoff et al. (1988) paper, which distinguishes between two and not three exit modes, is verified, as long as the coefficients to the explanatory variables are of interest. If one, on the other hand, focuses on the predictive ability of the models, the interesting conclusion is, that the simple financial distress model (where the exit to financial distress is modelled treating all other firms as censored) performs just as well as the parametric competing-risks model, and therefore, that if prediction is the sole purpose of the model, it does not matter which of these two models is estimated. The predictive performance of the pooled logit model (where all exits are pooled) is far worse than both of the other two models.

The paper is divided into 6 sections. The paper begins in section 2 with a discussion of economic theory and the identification of explanatory variables. Then, in section 3, data is described and the dependent variable is constructed. The theory underlying the estimation procedures as well as the results from the estimations are presented in section 4 and 5. Section 6 concludes.

2. Economic Theory and the Explanatory Variables

In this section the theoretical literature is presented and discussed, and various hypotheses on the factors influencing financial distress are set up. In section 4 the hypothesis will be tested and the results will be discussed. The explanatory variables have an effect on other exits (voluntary liquidations, mergers etc.) as well, but, as the focus is on the firms in financial distress, the discussions will take the point of departure in these firms.

For the most part empirical research in the area of bankruptcy prediction does not rest on any explicit theory, c.f. Beaver (1966), Altman (1968), Ohlson (1980), and Shumway (2001). The few empirical papers that do rest on explicit theory have a fragmented theoretical discussion. They do not present a full model "explaining" what drives firms into financial distress. Instead partial models each explaining some of the features of post-entry performance of firms, are typically presented. The same is done here. If nothing else is mentioned, the discussions are "all other things being equal" considerations. In the sample the variables are correlated.

The discussion is split up in three sections. The first section discusses what is called the "core indicators". These are the common indicators used in bankruptcy studies. They are indicators such as age, financial performance, and size. The second section discusses proxy variables, and the third section presents the controls used in the study.

In the appendix on data (section 7) all variables are listed, their definitions are given, and descriptive statistics are presented.

2.1. The Core Variables

In the following sub-sections some theoretical justifications for the inclusion of the core variables are given. Their expected effects on the probability of moving into financial distress are seen in table 2.1.

Table 2.1: Core variables and their expected effect

Variables	Expected effect on the probability of default
Firm Age (dummies)	Bell-shaped effect (see figure 2.1.1)
Short term debt to total assets	Increase
Return on net assets	Decrease
Solvency ratio	Decrease
Firm size	Decrease

2.1.1. Duration Dependence

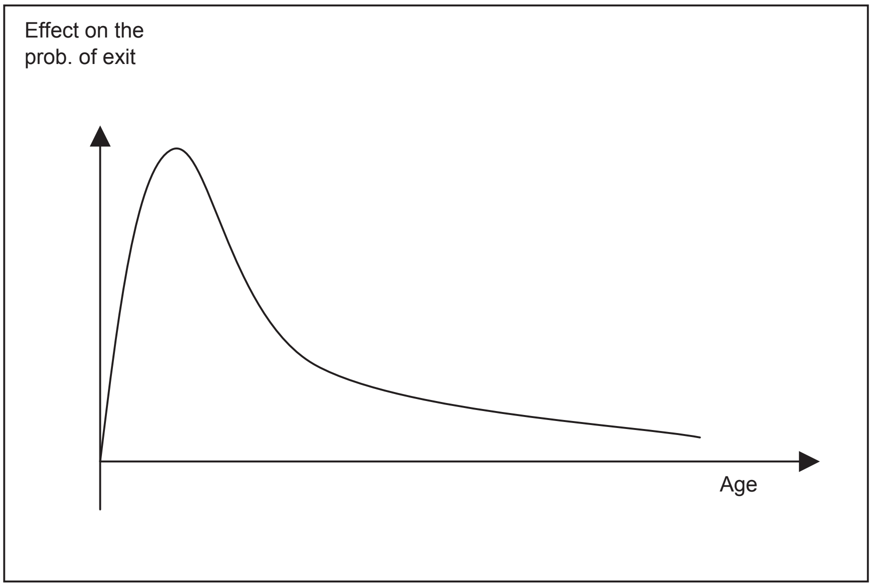
The effect of age (also called duration dependence) is of particular interest. The theory of Jovanovic (1982) and Pakes and Ericsson (1998) consider firm entry and firm exit. The theory suggests that the effect of age on firm exit is bell-shaped. In this section the mechanisms in their theoretical models are presented.

In the theoretical model in Jovanovic (1982), firms learn about their efficiency as they operate in the industry. Firms know the average market profitability, but they do not know their own potential. After entry they start to learn about their own profitability potential, and the firms either expand, contract or exit depending on where they are in the distribution of profitability. The efficient firms grow and survive, and the inefficient decline and fail.[2] Pakes and Ericsson (1998:39) show that many functional specifications of Jovanovic (1982)'s model imply that it takes time for entrant firms to acquire sufficient information about their parameters before they are able to decide whether they want to exit or to stay in the market. The implication of the model is that the effect of age on exits is bell-shaped, c.f. figure 2.1.1. When the firms are young they have not yet learned their own potential and the probability of exit is low. As time passes the firms learn about their own profitability potential, and the firms either expand, contract or exit. The hypothesis is that age has a bell-shaped effect.

---

[2] Jovanovic (1982)'s model is extended in Ericson and Pakes (1995). In Ericson and Pakes (1995) the firms are uncertain about the market's evaluation of the profitability of innovation. The firms enter the market and explore the economic environment actively and they invest to enhance productivity. The potential and actual productivity changes over time in response to effort and stochastic outcomes (of the firms own effort and effort of other firms in market). If successful the firms grow, otherwise they shrink or exit.

Figure 2.1.1: The effect of age on the probability of exit



2.1.2. Firm Performance

There is no consensus on which ratios should be used in a model that predicts firms that enter financial distress, but most studies include at least some measure of profitability, capital gearing and liquidity, c.f. table 2.1.2 which summarizes the predictors used in Beaver (1966), Altman (1968), Ohlson (1980) and Shumway (2001). In this paper the short-term debt to total assets, the companies' earnings capability (the return on net assets), and the solvency ratio are used.

A high debt ratio implies that companies may find it difficult to repay their debt. The hypothesis is that a high short-term debt to total assets increases the probability of moving into financial distress.

Return on net assets reflects the primary operating result as a ratio of the applied resources. A high return on net assets does not necessarily reflect that the company has a lower probability of entering financial distress. Instead, it might reflect that the firm takes high risk and is rewarded for it. As the legal status dummy and the location proxy measures the firms' willingness to take risk, the hypothesis is that a high return on net assets decreases the probability of moving into financial distress.

A high solvency ratio expresses the company's ability to generate satisfactory earnings over time, as rising profits are normally reflected in expansion of equity capital. The hypothesis is that a high solvency ratio decreases the probability of moving into financial distress.

Table 2.1.2: The predictor variables identified in the studies

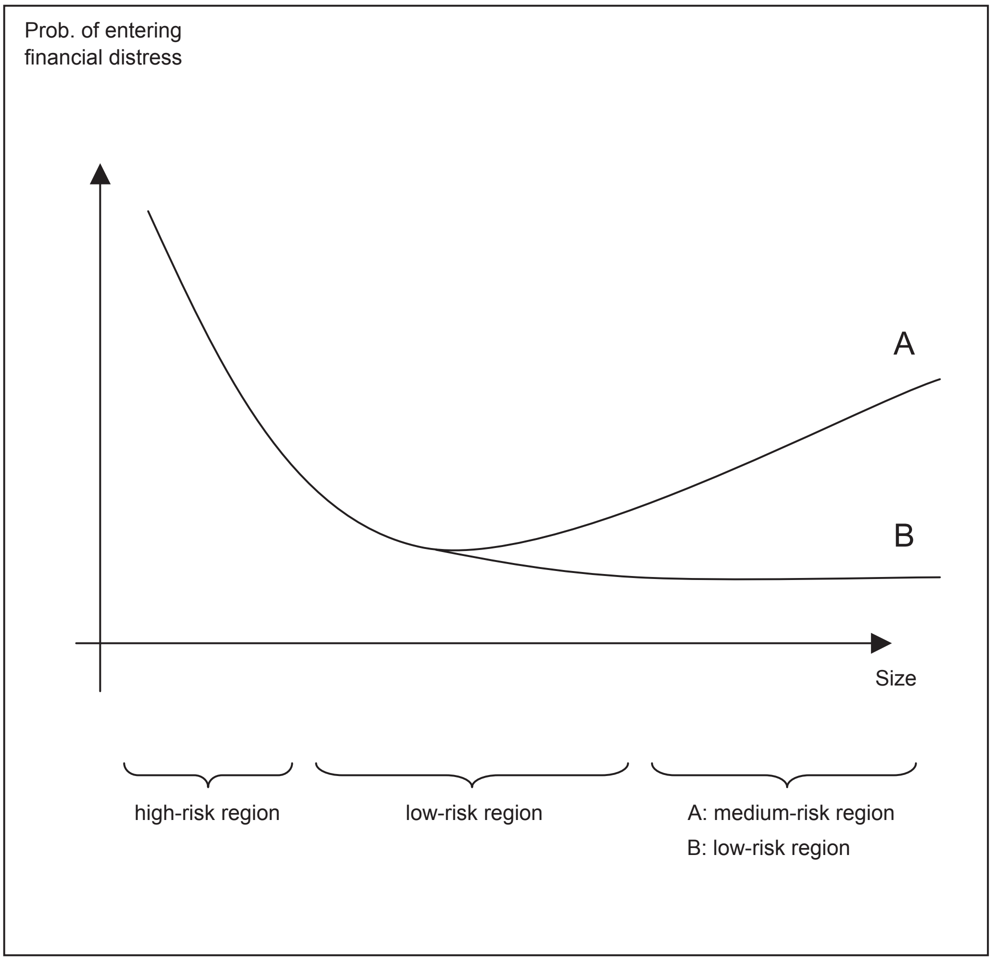
Beaver (1966)	Cash flow/total debt
Altman (1968)	Working capital/total assets Retained earnings/total assets Earnings before interest and taxes/total assets Sales/total sales Market value equity/book value of total debt
Ohlson (1980)	Log (total assets/GNP price-level index) Total liabilities/total assets Working capital/total assets Current liabilities/current assets A dummy = 1 if total liabilities exceeds total assets, 0 otherwise Net income/total assets Funds provided by operations/total liabilities A dummy = 1 if net income was negative for the last two years, 0 otherwise Change in net income
Shumway (2001)	Net income/total assets Total liabilities/total assets Market size Past stock returns The idiosyncratic standard deviation of stock returns

2.1.3. Firm Size

Firm size and firm age are expected to be correlated. In this section, the effect of size, given age, is discussed.

In figure 2.1.3 the effect of size on the probability of entering financial distress is sketched. Hypothesis A is that there exists an optimal firm size. This means that there is a trade-off between being relatively small and relatively large, and therefore that the effect of firm size on the probability of moving into financial distress is nearly U-shaped. The reasoning behind this hypothesis is that small firms have a higher probability of entering financial distress, because they are not so resistant to the shocks they might encounter, and that large firms have a high probability of entering financial distress, as they might have 1) inflexible organizations, 2) problems with monitoring managers and employees and 3) difficulties with providing efficient intra-firm communication.

Figure 2.1.3: The effect of size on the probability of entering financial distress



Hypothesis B is that the probability of entering financial distress decreases along with an increase in size. Hypothesis B is in line with the theoretical literature presented in box 2.1.3. As is discussed in the box, the theoretical models predict that the exit rates of the firms are a decreasing function of firm size, and so hypothesis B will be tested later on in section 4. In the estimations firm size is measured as log(total assets). The hypothesis is that an increase in firm size has a larger effect when the firm is relatively small, compared to the effect when the firm is relatively large.

Box 2.1.3: Studies on the effect of size

In the studies of Jovanovic and MacDonald (1994) and Klepper (1996), the focus is on firms that innovate. Both studies stress the superior ability of larger and older firms in order to adjust to drastic innovations. Jovanovic and MacDonald (1994) model a major (exogenous) technological change, which leads to exit of firms that are unable to innovate in the new regime. Klepper (1996) emphasises differences in firm innovative capabilities and the importance of firm size in appropriating the returns from innovation. The model has two simple forces. One is that the ability to appropriate the returns to process R&D depends centrally on the size of the firm. The other is that firms possess different types of expertise leading them to pursue different types of product innovations.

The ability to adapt to drastic innovations is closely related to the firms' access to the credit market. Brito and Mello (1995) analyse the problem of financing firms' production and opportunities when firms cannot secure sufficient internal funds and need additional external finance. There is asymmetric information between those that own and control the assets of the firm and outside investors. However, as time evolves, outside investors can learn more about the quality of the firms' management and accordingly adjust the terms of the financing contract. Brito and Mello (1995) show, that the size of the firm is correlated with the duration of the relationship between those who control the assets and those who finance the company. Their model implies that the exit rates are a decreasing function of firm size, which is consistent with the studies of Jovanovic and MacDonald (1994) and Klepper (1996). However, note that Frame, Srinivasan and Woosley (2001) find that credit scoring lowers information costs between borrowers and lenders, thereby reducing the value of traditional, local bank lending relationships.

2.2. Proxies

In the estimations in section 4 proxies are used for several of the variables that are inherently unobservable. In an estimation problem like this the use of proxies are preferred over a specification that leaves some of the variation to be modelled by an unobserved heterogeneity term, c.f. the discussion in Arellano (2003:11). The theory behind the use of proxies in a linear model is discussed in box 2.2.

The proxies used in this paper are summarized in table 2.2 and discussed in the following sections. Table 2.2 shows that motivation, uncertainty and ability are proxied by other variables. The relevant assumptions are assumed to hold, c.f. box 2.2.

Box 2.2: The use of proxies

The use of proxies is discussed in Wooldridge (2003:295ff). The idea is illustrated in a model with three independent variables and an error u : (1) $y = \beta_0 + \beta_1 x_1^* + \beta_2 x_2 + \beta_3 x_3 + u$.

In the model x_2 and x_3 are observed and x_1^* is not observed, but x_1 is a proxy for x_1^* . The proxy (x_1) is required to have some relationship with what it is a proxy for (x_1^*). In the standard case this is captured by the simple regression equation: (2) $x_1^* = \delta_0 + \delta_1 x_1 + v_1$, where v_1 is an error due to the fact that x_1^* and x_1 are not exactly related. The parameter δ_1 measures the relationship between x_1^* and x_1 . If δ_1 is equal to 0, then x_1 is not a suitable proxy for x_1^* . If δ_1 is different from 0, then x_1 is a suitable proxy for x_1^* . δ_0 is an intercept, which allows x_1^* and x_1 to be measured on different scales.

When estimating the model, proxies are used instead of the variable they are used as a proxy for. Wooldridge (2003:296) calls this the plug-in solution to the omitted variables problem. The assumptions needed for the method to provide consistent estimators of β_2 and β_3 are the following: The error u has to be uncorrelated with x_1^* , x_1 , x_2 and x_3 . The error v_1 is uncorrelated with x_1 , x_2 and x_3 .

2.2.1. Diversification

In the literature the effects of diversification on the *value of firms* (and not on the probability of entering financial distress) are discussed. In the model in Jovanovic (1993) the main reasons to diversify are gains in market power (firms with market power in two substitute product fields may be more profitable than two single product monopolies acting non-cooperatively), risk elimination, access to financial resources, and efficiency gains in production.

Rajan, Servaes and Zingales (2000) discuss different types of diversification. In the model firms can diversify in similar divisions and in divisions that differ from each other. The result is that diversification can be both value-enhancing and value-reducing. The model predicts that 1) if divisions are similar in the level of their resources and opportunities, funds will be transferred from divisions with poor opportunities to divisions with good opportunities, 2) when diversity in resources and opportunities increases, resources can flow towards the most inefficient investments and less value-able divisions. The model is tested and evidence consistent with it is found.

Table 2.2: Proxies and their expected effect

Variables	Expected effect on the probability of default	Proxy for
Diversification 2 sectors (related business) (dummy)	Decrease	Uncertainty
Diversification 3–9 sectors (related business) (dummy)	Decrease	Uncertainty
Diversification 2 sectors (unrelated business) (dummy)	Decrease	Uncertainty
Diversification 3–9 sectors (unrelated business) (dummy)	Decrease	Uncertainty
Local authority group 1 (reference dummy)		
Local authority group 2 (dummy)	?	Uncertainty/willingness to take on risk
Local authority group 3 (dummy)	?	Uncertainty/willingness to take on risk
Local authority group 4 (dummy)	?	Uncertainty/willingness to take on risk
Local authority group 5 (dummy)	?	Uncertainty/willingness to take on risk
Concentration	?	Uncertainty
Owned by the public (dummy)	?	Motivation
Owned by a fund (dummy)	?	Motivation
Ultimate parent companies (dummy)	?	Motivation
Wholly-owned subsidiaries (dummy)	Decrease	Motivation
Private limited liability company (dummy)	Increase	Motivation/willingness to take on risk
Public limited liability company (reference dummy)		
Publicly traded companies (dummy)	Decrease	Motivation
Critical comments from the auditors (dummy)	Increase	Ability

Some empirical studies which conclude that there is a loss in value when firms are diversified are Berger and Ofek (1995) and Lamont and Polk (2002). Berger and Ofek (1995) find that diversified firms have values that on average are 13 to 15 per cent below the sum of the imputed stand-alone values of their segments (measured as if they were operated as separate firms), and that the loss in value is considerably less for related diversifications. The loss in value can have two explanations: Diversification itself might somehow destroy value, and diversification and lower value may not be causally related, but instead reflects firms' endogenous choices (e.g. low value firms choose to diversify, leading to a negative correlation between diversification and value). Lamont and Polk (2002) take the analysis of Berger and Ofek (1995) a step further in order to identify whether there is a causal relation or not. They find that exogenous changes in diversity, due to changes in industry investment, are negatively related to firm value, and thus that diversification destroys value.

The conclusion from the theoretical and the empirical studies is that diversification can have both value-enhancing and value-reducing effects. Nonetheless, in the context of financial distress, given that there is already controlled for a high/low return on net assets, diversification is considered only positive. Diversification is an (sometimes expensive) insurance against economic shocks etc., and so the hypothesis is that diversified firms have a lower probability of entering financial distress.

2.2.2. The Location of the Firm

The location of a firm is used as a proxy for uncertainty as well as for the willingness to take on risk, as the location of a firm reflects how the uncertainty and the willingness to take on risk is perceived. Factors influencing the choice of location could be the availability of labour and the location of competitors, customers and potential partners, c.f. Tirole (1997:277ff), who among other things discusses the so-called location or spatial-differentiation model in which different consumers are located at different places. No hypothesis is set up on the effect of the location of the firm. It is left to the estimations to show the effects.

2.2.3. Concentration

The competitiveness in a specific industry is difficult to measure unless one observes prices in industries with similar cost structures, the temporal pattern of the industry price or one measures accurately the firms' marginal cost (Tirole (1997:221)). This type of information is often difficult to obtain. Information that is easier to obtain is information of the rates of profit and the firms' market shares. Firms' market shares can be used to construct concentration indexes. Several widely used indexes exist, e.g. the m-firm concentration ratio, the Herfindahl index and the entropy index (Tirole (1997:221f)).

There is no clear link between concentration indices and profitability, however, it is pointed out in Tirole (1997:222), that most cross-sectional analyses find weak but statistically significant link between concentration and profitability. Tirole (1997:223) concludes that concentration indices are useful as they give an easily computable and interpretable indication of how competitive the industry is. Nonetheless, it is important to note that a high (low) degree of concentration does not necessarily lead to low (high) competition. This is discussed by the Danish Competition Authority in their report called “Konkurrenceredegørelsen 2003” (in Danish), c.f. Konkurrencestyrelsen (2003:chapter 2.3). The Danish Competition Authority points out that one could think of situations where there is a high degree of competition, even though the concentration is high, and that one could also think of situations where there is a low degree of competition even though there are many companies in the market, e.g. if there is a cartel.

As a proxy for the competitiveness of an industry, and thereby the uncertainty that the firm is facing, the concentration index is used in the estimations. The concentration index is measured as the sum of the market shares in the four largest companies as a percentage of the total domestic turnover in a specific sector.[3] No hypothesis is set up on the effect of the concentration variable as this discussion shows that the effects are not necessarily clear.

2.2.4. Public/fund Ownership

Ownership is used as a proxy for motivation. A small number of firms are owned by the public sector or owned by a fund. If the public sector and/or the fund wanted to inject funds when needed, none of these firms will move into financial distress. The question is whether one believes that the public sector and/or funds are prepared to do so. The arguments can go either way. Even though one might think that the public sector or a fund would be more prepared to do so, no hypothesis on the effect of ownership is set up. It is left to the estimations to show whether there is a significant effect of these ownerships variables at all (a dummy for having the public sector as an owner and a dummy for having a fund as an owner).

2.2.5. Critical Comments from the Auditors

Numerous studies have shown that human capital endowment of the entrepreneur is an important determinant for firm performance, e.g. Bates (1990) and Statistics Denmark (2001). Here ability is proxied by a firm having attached a critical comment from the auditors to its financial statement. The critical comments included in the “critical comments from the auditors”-dummy do not explicitly express that the company is about to fail. The comments are argued to point towards the inability of the entrepreneur. (An alternative interpretation could be that it reflects uncertainty concerning the true value of the company). The comments included are for example the following: “illegal loans have been adopted”, there are “inconsistencies in the profit and loss account” or “the financial statement is incomplete”. Critical comments from the auditors give a warning signal to creditors, and the hypothesis is that firms with critical comments from the auditors have a higher probability of moving into financial distress.

2.2.6. Wholly Owned Subsidiaries and Ultimate Parent Companies

The creation of a group of companies results in the ultimate parent company being able to control all issues concerning its subsidiaries, e.g. the way various companies undertake internal transactions. Since the transactions between the companies in a group of companies have an effect on the financial statements of all the involved companies, the consequences of internal transactions are relevant when analysing the accounts of the various companies belonging to the same

---

[3] Note that the concentration index suffers from the fact that it considers the economy as a closed economy.

group. In the same way, when analysing liquidity in a group of companies, the financial statements of the various companies belonging to the group must be taken into account. As the financial conditions of an ultimate parent company and its subsidiaries are intertwined, it is important to incorporate information on ultimate parent companies and subsidiaries in a model that identifies firms in financial distress.

Audretsch and Mahmood (1995:100), who analyse new firms' survival, argue: "Because an established firm already has experience about the specific economic conditions and managerial competence, when a branch or subsidiary plant is opened it should face a systematically lower likelihood of failure. That is, the hazard rate would be expected to be systematically greater for new establishments which are independent enterprises and systematically lower for new branch or subsidiary plants opened by an incumbent enterprise." Audretsch and Mahmood (1995:102) do find that the hazard rate tends to be greater for new firms than for new branch plants opened by existing enterprises. In line with this result, the hypothesis is that wholly owned subsidiaries have a lower probability of moving into financial distress compared to companies that are not wholly owned subsidiaries.

It is not clear whether ultimate parent companies have a higher or lower probability of moving into financial distress. One could argue that ultimate parent companies have an advantage compared to other companies, as they are able to control all issues concerning their subsidiaries. On the other hand, ultimate parent companies might have a tendency to let the not-so-well-performing subsidiaries drain resources from the group as a whole, including the ultimate parent company, e.g. because of reputational risk. This would imply that ultimate parent companies would have a larger probability of entering financial distress compared to other companies. It is left to the estimations to show the effect of being an ultimate parent company.

2.2.7. Limited Liability

Public limited liability and private limited liability companies are analyzed. Public limited liability companies and private limited liability companies need at least 500,000 and 125,000 Danish kr., respectively, as share capital, when they are set up, and the owner of one of these two types of companies is not liable for more than the amount of the share capital.[4] The effect of limited liability is discussed in Stiglitz and Weiss (1981) and estimated in Harhoff et al. (1998). Stiglitz and Weiss (1981) model credit rationing in markets with imperfect information. Assuming that

---

4 Note that not all public limited liability companies are listed on a stock exchange and that by law private limited liability companies cannot be listed on a stock exchange. The terms public and private refer to the “type” of company (i.e. the law the specific company is following).

the projects are undertaken by the firms in their model are the sole projects, and that there is limited liability, Stiglitz and Weiss (1981) show that in equilibrium a loan market may be characterized by credit rationing. Credit rationing is defined in Stiglitz and Weiss (1981) as circumstances in which either a) among loan applicants who appear to be identical some receive loan and others do not, and the rejected participants would not receive a loan even if they would offer to pay a higher interest rate or b) there are identifiable groups of individuals in the population who, with a given supply of credit, are unable to obtain loans at any interest rate, even though with a larger supply of credit they would. In the equilibrium the entrepreneurs choose projects characterized by a relatively high, expected return and a relatively high risk of failure. Furthermore, Stiglitz and Weiss (1981) show that limited liability increases the risk the entrepreneurs are taking. Building on the theory of Stiglitz and Weiss (1981), Harhoff et al. (1998:459ff) test the hypothesis that firms with limited liability will experience a comparatively high risk of insolvency relative to firms operating under full liability. Harhoff et al. (1998) do find that firms under limited liability have significantly higher failure rates than sole proprietorships. As private limited liability companies need less share capital than public limited liability companies, following Stiglitz and Weiss (1981) and Harhoff et al. (1998), the hypothesis is that the probability of moving into financial distress is higher for private limited liability companies.

2.2.8. Companies Listed on a Stock Exchange

By law private limited liability companies cannot go public by being quoted on a stock exchange, whereas public limited liability companies can. There may be many motives behind a company's decision to be listed. The primary reason for the majority of companies is access to the capital market. Harhoff et al. (1998:470) find that stock-based corporate firms ("Aktiengesellschaft and Kommanditgesellschaft auf Aktien") have a relatively low failure risk, and they argue that this might reflect that (in their sample) many stock companies are listed, and that they therefore have been subject to thorough screening by banks before their shares are traded. The hypothesis is that the same is true for Danish non-financial firms, in other words, that the public limited liability companies that are listed have a lower probability of moving into financial distress compared to companies that are not listed.

2.3. Controls

In the estimations controls are set in, c.f. table 2.3. Below the various controls are discussed in turn.

Table 2.3: Controls used in the estimations

Variables	Control/expected effect on the probability of default
Macroeconomic environment (Year dummies): Year 1996 (reference dummy), Year 1997, Year 1998, Year 1999, Year 2000, Year 2001	Included to control for the macroeconomic environment
Sector Affiliation dummies: Farming, Forestry, Fishing, Mining, Manufacturing (reference dummy), Energy, Construction, Trade and hotel, Transport, Business service, Public service activities, Organisations, Not stated, Unknown	Included to control for sector affiliation
IT dummy	Increase
Some firms register a Primary Bank connection in one of the following four categories (see the appendix on data, section 7, for further details): Category 1 (dummy), Category 2 (dummy), Category 3 (dummy), Category 4 (dummy), Firm's that do not register a primary bank connection (reference dummy)	?

Note: There is no year dummy in 1995 as, by construction, no firms leave the database between 1995 and 1996.

2.3.1. Sector Affiliation

It would be natural to hypothesize that the probability of entering financial distress depends on the sector affiliation of the firm. Here theoretical and empirical literature underpinning this statement is reviewed.

In a theoretical model Gort and Klepper (1982) show that technological and knowledge conditions determine the relative ease with which new firms are able to innovate and therefore survive. Audretch (1991), who studies the survival rates at the industry level, tests the model and finds that survival rates vary considerably across industries, and that they are shaped by the conditions of technology and demand underlying the industry. Audretch and Mahmood (1995) take the analysis a step further and identify explicitly the post-entry performance of new businesses by linking their likelihood of survival to the conditions of technology and demand underlying the industry within which they operate. They find that the likelihood of a new business surviving is in fact shaped by the underlying technological conditions and extent of scale economics among other things. The empirical evidence in Agarwal and Audretsch (2001) also suggests that the relationship between firm size and the likelihood of survival is shaped by technology and the stage of the industry life cycle.

The hypothesis is that the probability of moving into financial distress varies between different business sectors, and so dummies for sector affiliation are used in the estimations. The division of companies into the various sectors follows the NACE classification. As there is no NACE sector affiliation category called "IT and telecommunication companies" a dummy variable is constructed after the other dummies are specified. For details of how the dummies are constructed the reader is referred to the appendix on data (section 7).

2.3.2. Macroeconomic Effects

Several studies have found that business cycle effects influence movements in and out of financial distress, e.g. Audretsch and Mahmood (1995), who among other things examine the link between the business cycle and the exposure to risk by including the unemployment rate. They find a positive coefficient of the unemployment rate, which suggests that the hazard rate for new establishments tends to be greater during periods of higher unemployment (which is commonly associated with macroeconomic downturns). In line with Audretsch and Mahmood (1995) our hypothesis is that macroeconomic downturns increase the likelihood of moving into financial distress. In the estimations, year dummies are included to control for the macroeconomic environment. They affect all the firms in the same way and they serve as a robust control for common aggregate effects. Any aggregate variable will be a linear combination of the time dummies, and thus, if a full set of time dummies is included any aggregate variable will be perfectly co-linear with them and hence redundant, as is discussed in Arellano (2003:61). Note that in this paper only estimations of the aggregate process are estimated. The actual process is not specified. This is the best that can be done with a small T panel.

2.3.3. Firms with a Primary Bank

Some firms register their primary bank connection. If the fact that the firms register a primary bank can be thought of as relationship lending, then the fact that a firm registers a bank under “primary bank” can be interpreted as a more outspoken banking relationship compared to the relationships of the firms that do not register a primary bank. In line with the paper of Brito and Mello (1995), presented in box 2.1.3, the hypothesis would be that the firms that register a primary bank have a lower probability of moving into financial distress. Here, no hypothesis is set up. It is not clear whether there is an effect of a registration of a primary bank on the probability of moving into financial distress. In the estimations, the banks are split up according to the four categories 1, 2, 3 and 4, based on the volume of working capital of the banks (working capital comprises deposits, issued bonds, subordinate capital, and equity capital), c.f. the data description in the appendix on data (section 7).

2.4. Conclusion

The discussion in this section suggests that the probability of moving into financial distress depends on the variables listed in table 2.1. Table 2.2 lists the proxies. Table 2.3 lists the controls used in the estimations.

In the estimations in section 4 all explanatory variables are treated as strictly exogenous variables, that is, the information on the firms is taken as given and uncorrelated with unobservables. The exogeneity assumption is perhaps more reasonable here than in most cases due to the fact that the model is estimated on a rich data set, and so there are several proxies for the variables that are inherently unobservable. Proxies are important. Take the example of a potential endogenous variable: the solvency ratio, which is calculated as equity capital over total assets. As the firms have some command over the level of equity capital (e.g. how much they pay to their shareholders), whether they pay more or less to their shareholders might very well be correlated with the degree of uncertainty that characterizes the economic environment they face. This is usually a problem as uncertainty is not included in the estimations. Here, it is less of a problem simply because the four diversification variables as well as the location dummies and the concentration index are used as proxies for the uncertainty that the firms are facing.

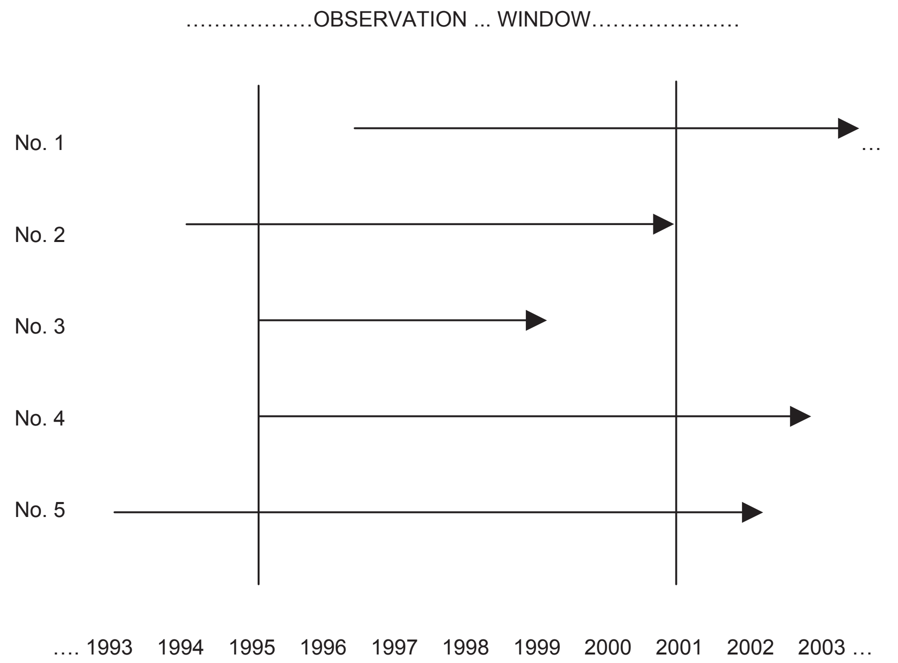
3. Data and the Construction of the Dependent Variable

In this section data is presented, and the construction of the dependent variable is discussed. In the first section an introduction to the database is given, the sample selection procedure is presented, and special issues concerning duration data is discussed. Then a thorough discussion of the construction of the dependent variable follows, and an overview of the data set is given.

3.1. The Data Base and Sample Selection

Data is obtained from KOB A/S. The data base covers all Danish public limited liability companies (“aktieselskaber”) and private limited liability companies (“anpartsselskaber”) that existed in the period from 1995 to 2001. The data set consists of a *single spell* for each firm. Once a firm has exited it cannot re-enter. This type of data is called duration data. Duration data can be either flow or stock sampled (Wooldridge (2002:chapter 20)). Flow sampling occurs when individuals or firms, that enter the state of interest at some point during an interval, are sampled, whereas stock sampling occurs when only individuals or firms, that are in the initial state at a given point in time, are sampled. The present data is a combination of stock- and flow-sampled data. The firms that were incorporated before 1995 but only observed since 1995 are stock sampled, since the firms are only observed between 1995 and 2001 no matter when they were incorporated. The new firms entering between 1995 and 2001 are flow sampled. An event is censored if the exact time at which the event occurs is not observed. Both flow- and stock-sampled data can be right censored, whereas only stock-sampled data can be left censored or left truncated, see D'Addio and Rosholm (2002: 4ff).[5] In the data set, if the firms have not exited before 2001, it is right censored. There is no left censoring as the incorporation date of all companies is known, however, the stock sampled spells are left truncated as the spells are in progress when the observation period starts. The flow sampled spells are not left truncated as they start at some point "in the observation window", c.f. box 3.1.

Box 3.1: 5 different firms (examples of spells)



Five different representative firms from the dataset are sketched. Spell no. 2 and 5 represent stock sampled spells, whereas spell no. 1, 3 and 4 represent flow sampled spells. The firms in the box are censored in various ways. Firm no. 3 is not censored at all. The spell is complete as the full length of the spell is observed. Firm no. 4 is right censored as the firm is not observed beyond 2001. It is neither left censored nor left truncated as the firm is observed from the incorporation date. Firm no. 5 would have been left censored had the starting data not been known. However, the starting date is known and the firm is left truncated. It is also right censored. Firm no. 2 is left truncated and not right censored (as the spell ends in 2001). Firm no. 1 is right censored in 2001.

---

5 Right censoring refers to the situation where there is missing information on the times and states occupied after the end of a given observation period. Left censoring refers to the situation where a spell is in progress at the beginning of the observation period and where only the duration from that point in time is observed. Left truncation refers to a situation where a spell is in progress at the beginning of the observation period and where the duration from that point in time is observed and where at the same time the origin date is known. The important difference between left censoring and left truncation is that when a spell is left censored, the origin date of the spell in progress at the start of the observation period is *unknown*, whereas when a spell is left truncated the origin date of the spell in progress at the start of the observation period is *known.*

The data base comprises information on financial issues as well as non-financial issues. KOB A/S receives annually accounts data and information on the status of the companies from the Danish Commerce and Companies Agency. The companies are classified as active and inactive firms, respectively. As will be discussed in section 3.2, the inactive firms can be firms that are financially distressed (e.g. firms that have gone bankrupt), voluntarily liquidated firms, or firms that have merged with other firms, etc. At least once a year, KOB A/S conducts a telephone interview with each of the companies to confirm and supplement the accounts data with information on number of employees, sector affiliation, etc. On top of this information received from KOB A/S, new information is incorporated in the data base. The data base is augmented to also include whether or not the company is 1) an ultimate parent company, 2) a wholly owned subsidiary, 3) quoted on the stock exchange, 4) owned by the public, 5) owned by a fund and 6) a concentration index (measuring the concentration of the various sectors).[6]

In the raw data set there are 603,956 firm-year observations covering the period 1995-2001. After the exclusion of holding companies and financial firms and after making some corrections to the data base[7], there are 430,422 firm-year observations left. Firms with illogical variables are excluded (such as short-term debt less than zero and a solvency ratio larger than 100 per cent) and remaining are 409,906 firm-year observations. A panel is then constructed.

The panel comprises:

1) companies incorporated in the period 1995-2001 with at least 5 employees the year they are included in the sample and with a balance sheet of at least kr. 500,000 (the flow sampled companies),

2) companies that were active in 1995 but were incorporated before 1995 (the stock sampled companies) with a balance sheet of at least kr. 500,000 and 5 employees in 1995.

---

[6] Hans Christian Kongsted and Kasper Nielsen have helped me with information on items 1), 2), 4) and 5). The Danish Competition Authority has kindly provided the concentration index.

[7] The *financial statement* for a specific firm is deleted if it is not the last financial statement for the company in a specific year in order to insure that there is only one financial statement from each company each year. Furthermore, *the company* is deleted from the database, if 1) it has had a net return on assets above 300 pct. (or below -300 per cent) in one specific year, if 2) the company is financially distressed in 1995 (so that there is not a financial statement from it when it can be classified as an active company), if 3) the company does not hand in a financial statement each year, if 4) the company is assessed not to be relevant for the analysis, e.g. A/S Storebæltsforbindelsen, which is a government-guaranteed entity. Government-guaranteed entities are typically structured as government-owned limited-liability companies. Their tasks are defined in an act or legal document, which gives access to government guarantees for loans within a certain limit, c.f. Danmarks Nationalbank (2003:chapter 7.5).

After the application of the criteria the panel consists of 168,778 firm-year observations, covering 32,453 firms. Due to missing variables, the final number of firm-year observations in the estimations is 168,350, covering 32,365 firms.[8]

By only including firms with 5 employees or more and a balance sheet of kr. 500,000 or more in the dataset, it is made sure, that only active firms are analyzed. On the basis of the balance sheet criteria kr. 500.000 some of the newly started private limited liability companies might be excluded from the analysis (kr. 125.000 are needed as share capital to start a private limited liability company), whereas newly started public limited liability companies are included, if they still have the start-up capital of kr. 500.000, which are required to set up a public limited liability company. As described above the firms enter the database when they have kr. 500.000 on the balance sheet and 5 employees. Once they are included in the data set, they do not get excluded if they obtain a balance sheet below the limit or if they have less than 5 employees.

Comparisons of the data used in this paper with the data used in some of the studies mentioned in the introduction are seen in tables 3.1.a and 3.1.b. Table 3.1.a sketches the time period covered by the other studies as well as the number of observations included in the studies. The time periods covered and the "amount" of information differs quite a lot from study to study. Compared to the other studies, the data for this study is unique as it comprises information on the whole population of Danish firms, altogether more than 150,000 firm-year observations (of which 2,617 firms enter financial distress, c.f. section 3.2 which discusses the definition of distress in detail). Table 3.1.b summarizes the type of data used in the different studies. Looking at the table, several issues are to be noted right away. First of all, it is seen that the analyses of Beaver (1966) and Altman (1968) are based on paired sample design. They might have a sample selection problem due to the “matching” procedures. As is discussed in Ohlson (1980) it is not known what is really gained or lost by different matching procedures, including no matching at all. Secondly, all companies analysed in the different studies, except in this one, are publicly held companies. Companies listed on a stock exchange tend to be larger than companies that are not listed on a stock exchange. In this study the focus is on Danish firms. As most Danish firms are small and medium-sized enterprises (SME), not only large firms are considered. Thirdly, in the other studies, all analysed companies are industrials. Here firms from all the sectors in the economy are covered.

---

[8] Even though the companies are included in the database when they meet the above criteria, their age is calculated from the incorporation date. For the firms incorporated in the period 1995-2001, one could count the firm age, as of the date these firms enter the sample, however, this cannot be done for the firms that are already active in 1995. It is chosen to treat the firms in the same way, and so the age is calculated from the incorporation date.

Table 3.1.a: Time period and number of observations

Beaver (1966)	1954-1964	79 non-failed firms and 79 failed firms (59 were bankrupt, 16 involved non payment of preferred stock dividends, 3 were bond defaults, and 1 was an overdrawn bank account)
Altman (1968)	1946-1965	33 non-bankrupt and 33 bankrupt firms
Ohlson (1980)	1970-1976	2,058 non-bankrupt and 105 bankrupt firms
Shumway (2001)	1962-1992	28,664 firm years and 239 bankruptcies
This paper	1995-2001	More than 30,000 firms and more than 150,000 firm-year observations. 2,617 firms in financial distress, 907 voluntarily liquidated firms, and 1,233 firms that are acquired/have merged with other firms or the like.

Table 3.1.b: Data used for the studies

Beaver (1966)	Paired sample design: For every failed firm in the sample, there is a non-failed firm from the same industry and from approximately the same asset size class. Listed industrials.
Altman (1968)	Paired sample design: For every failed firm in the sample, there is a non-failed firm from the same industry and from approximately the same asset size class. Listed manufacturing corporations.
Ohlson (1980)	Industrials with traded equity (traded on some stock exchange or over-the-counter market).
Shumway (2001)	Industrials with traded equity (traded on the New York Stock Exchange and American Stock Exchange are included in the sample).
This paper	The whole population of public and private limited liability companies (mostly SMEs), covering all sectors of the Danish economy (except the financial sector).

3.2. The Construction of the Dependent Variable

The focus in this paper is on financially distressed firms. Some of the terms that are often used in the literature on corporate distress are failure, insolvency, default and bankruptcy. They are sometimes used interchangeably, although formally each of them can be defined in a different way, e.g. failure can be defined as the inability of a business to continue, especially through lack of money, insolvency indicates that the real net worth of the firm is negative, default refers to failure to do something that is demanded (e.g. failure to fulfil a contract, such as paying one's debts), and bankruptcy refers to a firm's formal declaration of bankruptcy. The definitions used in some of the studies mentioned in the introduction are summed up in table 3.2.a. Most of the studies use a purely legalistic definition of bankruptcy. After the presentation and discussion of the various exit codes in the data base, the definition of financial distress used in this paper is presented and discussed.

Table 3.2.a: The definition of failure/bankruptcy

Beaver (1966)	Failure – defined as the inability of a firm to pay its financial obligations as they mature. Operationally a firm is said to have failed when any of the following events have occurred: bankruptcy, bond default, an overdrawn bank account, or non-payment of a preferred stock dividend.
Altman (1968)	Bankruptcy is defined as those firms that are legally bankrupt and either placed in receivership or have been granted the right to reorganize under the provision of the National Bankruptcy Act.
Ohlson (1980)	The definition of failed firms is purely legalistic. The failed firms must have filed for bankruptcy in the sense of Chapter X, Chapter XI, or some other notification indicating bankruptcy proceedings.
Shumway (2001)	The definition of bankruptcy: Firms that filed for any type of bankruptcy are considered bankrupt.
This paper	The definition of firms in financial distress: (inactive) firms that have gone bankrupt, firms that have been compulsorily wound up and (active) firms that have experienced a write down of their debt or a forced sale.

In the data base, the companies are split up on the ones that are active and inactive. The inactive firms can exit the data base for various reasons: They can go bankrupt, they can be compulsorily wound up (“tvangsopløst”), they can be voluntarily liquidated (“likvidation”) and they can merge, etc. with other companies. The basis for these first three registrations in the database is the registrations in the official documents “Registreringstidende” and “Statstidende”. Whether or not the company has merged with other companies, or it has exited for other reasons, is registered by KOB A/S. The active firms are registered if they have had their debt written down (“tvangsakkord”) or if they have experienced a forced sale (“tvangsauktion”). By construction, the active firms that have experienced a write-down of their debt or a forced sale do not appear in the database after these events.

The registrations of the firms that have gone bankrupt, the firms that have been compulsorily wound up and the firms that have been voluntarily liquidated are the responsibility of the Danish Ministry of Economic and Business Affairs. However, in practice, the Danish Commerce and Companies Agency, which is under the auspices of the Danish Ministry of Economic and Business affairs, takes care of the administration of company and enterprise legislation, the registration and disclosure of certain information and documents about companies, including company accounts, etc. A firm has a *bankruptcy code* if it has formally been declared or if it is in the process of being declared bankrupt. To be *compulsorily wound up* means, that the Danish Commerce and Companies Agency has closed or decided to close the company. This can happen in situations where a company does not meet the legal requirements, e.g. when a company has not handed in its financial statement. Often it is financially distressed firms that do not fulfil the legal requirements.[9] A firm is *voluntarily liquidated* when the owner of a firm decides to close down the business. In this situation the creditors are paid, and the firm is then closed. This could happen in family-owned firms, when there is no generation to take over. It is central to note, that firms that are voluntarily liquidated do not inflict losses on their creditors. See also box 3.2.a for details on a firm's last financial statement and the registration of the codes.

Box 3.2.a: A firm's last financial statement and the registration of the codes

active	active	no inform.	bankruptcy code	: Raw data
active	bankrupt	"real bankruptcy" recorded as the day after the firm hands in its last financial statement	N.A.	: Data set
1995	1996	no inform.	14th September 1997	
fin. statem.	fin. statem.	no inform.	no inform.	

There is a lag between a firm's last financial statement and the registration of the codes (bankruptcy, compulsorily winding-up, voluntarily liquidated etc.). Take the example of a bankrupt firm. The timing of events could follow the time line sketched above. According to the raw data, the firm is active in 1995 and in 1996, and it hands in a financial statement both years. In September 1997 the firm gets a bankruptcy code from the Danish Commerce and Companies Agency. Apart from the code there is no further information on the firm in September 1997. The amount of time that passes by between the company hands in its last financial statement in 1996 and till it gets, in this case, the bankruptcy code, is arbitrary. It depends on the bankruptcy court that handles the specific case (how many cases it has already, etc). For this reason the data set is constructed so that the day after the last financial statement is handed in is recorded as the "actual" time of the bankruptcy.

---

[9] Møller et al. (1998:318f).

Box 3.2.b: Types of exits

E1 (= financially distressed firms)

E2 (= voluntarily liquidated firms)

E3 (= firms that have merged etc.)

All exits = E1 + E2 + E3

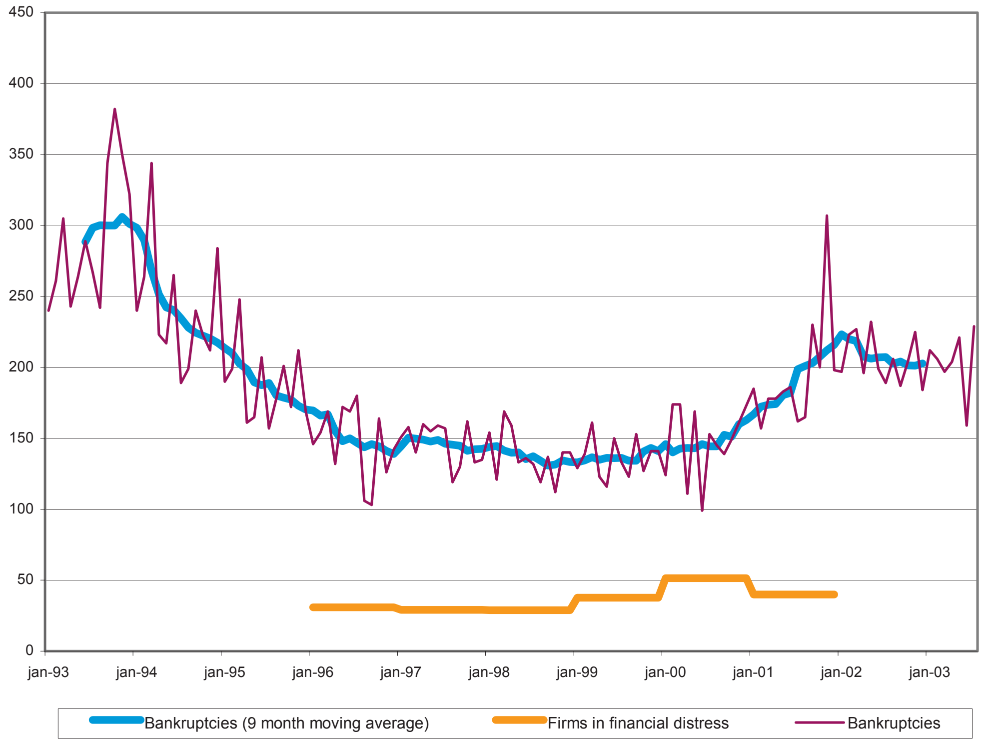
Table 3.2.b: Number of firms

	E1	E2	E3	Active	Total
1995	0	0	0	18853	28853
1996	372	87	177	20684	21320
1997	348	110	156	22008	22622
1998	347	129	195	23422	24093
1999	453	124	211	25000	25788
2000	618	148	226	26415	27407
2001	479	309	268	27639	28695
Total	**2617**	**907**	**1233**	**164021**	**168778**

Note: The number of firms that exit because of financial distress is 479 in 2001. This figure might be a little lower than the "actual" number. As was mentioned above, it takes time before the firms get the codes from the Danish Commerce and Companies Agency. The last information on the firms in the data set was incorporated in spring 2003. The data set covers firms that have handed in their financial statement in 2001 at the latest. If a potential bankrupt firm has not yet got the bankruptcy code in spring 2003 (but might have got it in fall 2003), it will not be recorded as a bankrupt firm in the data set. There could be a few firms for which this would be the case, and therefore the "actual" number of firms entering financial distress in 2001 is probably higher than the recorded number 479.

The focus in this paper is on the firms that end up in financial distress, or in other words, firms that can be expected to inflict a loss on the financial sector. In the measure of financially distressed firms the following firms are included: (inactive) firms that have gone bankrupt, firms that have been compulsorily wound up and (active) firms that have experienced a write down of their debt or a forced sale. The firms in financial distress are referred to as E1 firms. The firms that leave the data base for other reasons include firms that have been voluntarily liquidated (E2 firms) and firms that have exited because of a merger or acquisition (E3 firms). By construction, all exits are equal to E1+E2+E3, c.f. box 3.2.b.

Figure 3.2: Bankruptcies and firms in financial distress



Source: Own calculations based on the constructed data set and Statistics Denmark
Note: In the data base the firms in financial distress are registered on a yearly basis. The number of bankruptcies is recorded at a monthly basis at Statistics Denmark. In order to be able to include the numbers in the same figure in a meaningful way, the yearly number of firms in financial distress (in the data set) is pictured as the average monthly number of firms that enters financial distress.

Table 3.2.b shows the number of active firms every year and the number of firms that exit for E1, E2 and E3 reasons. The change in number of active firms is not equal to the number of exits as new firms enter the data base every year. Most of the firms that exit the data base, exit because of financial distress. Compulsorily liquidations account for the smallest number of exits.

The trend in the number of firms that leave the data set follows the trend in the number of official bankruptcies in the period 1995-2001, c.f. figure 3.2.[10] The official number of bankruptcies before 1995 was at a high level. When using the estimated model for prediction purposes, one should think of the consequence of the fact that the parameter estimates are obtained using a sample where a large number of the weakest firms have already exited. Underlying this observation is of course the

10 The differences between the official number of bankruptcies (obtained from Statistics Denmark) and the number of firms in financial distress (according to the data base) are due to the fact that there is 1) no sample selection behind the figures that Statistics Denmark reports, 2) the timing issues, c.f. box 3.2.a and 3) due to the fact that the definition of firms in financial distress in the constructed data set differs from the definition of bankruptcies.

assumption that the official number of bankruptcies is an indicator for the number of firms in financial distress. Furthermore, note that the sample period is relatively peaceful, and that the dataset does not cover a whole business cycle (as data is not available before 1995).

4. Econometric Theory behind and Estimation of a Competing-risks Model

This section presents and discusses the econometric theory behind the estimations, and the estimations results are reported. The suggestion is to estimate a parametric competing-risks model. As will be clear from the discussions in section 4.1, when estimating a parametric competing-risks model, one may proceed in several ways, depending on the data available. The strategy here is to follow Allison (1982) and therefore to assume independent exits and a special form of the destination-specific hazards. In section 4.2 the results from the estimations are presented. This section builds mainly on Allison (1982), Kiefer (1988), Jenkins (1995), Wooldridge (2002), Greene (2003), and Jenkins (2003).

4.1. Econometric Theory

The firms leave the data base for E1, E2 and E3 reasons. The occurrence of one type of event removes the firm at risk of all other event types, and so the method of competing-risks models is to be considered. Formally competing-risks models can be characterized in the following way, c.f. Crowder (2001): Let E1(t) be the hazard rate of exit to financial distress and the latent failure time is T_1, E2(t) be the hazard rate of exit to voluntary liquidation and the latent failure time is T_2, and E3(t) be the hazard rate of exit to mergers, acquisitions etc. with the latent failure time T_3. It is observed in the data whether there is no event at all (a censored case corresponding to an active firm, with spell length T_4), or whether there is an E1, E2 or E3 exit. Therefore the observed failure time is $timeT = \min\{T_1, T_2, T_3, T_4\}$.

When estimating a competing-risks model one can follow several estimations strategies depending on the data at hand, c.f. Jenkins (2003). With continuous data and the assumption of independent exits each destination-specific hazard rate can be thought of as the hazard rate that would apply were transitions to all the other destinations not possible, and one can maximize the overall log likelihood by maximizing the component parts separately. In practice, this can be done very easily. One should simply define new destination-specific censoring variables and then estimate separate models for each destination state.

Table 4.1: Grouped duration data and the five main assumptions in the literature

Main assumption	Main implications
Transitions can only occur at the boundaries of the intervals	Same result as in the continuous case (where the overall independent competing risk model can be estimated by estimating separate destination-specific models).
Destination-specific density functions are constant within each interval (though may vary between intervals)	When the destination-specific hazards become infinitesimally small, the likelihood contributions tend to expressions that are the same as in the above case.
Destinations specific hazard rates are constant within each interval (though may vary between intervals)	The expression for the likelihood contribution is not separable into destination-specific competing-risk models.
The hazard rate takes a particular proportional hazards form	This might be thought of as a more flexible specification of the hazard function. However note that it cannot be identified without further assumptions.
The log of the integrated hazard changes linearly over the interval	The assumption implies that the hazard increases within each interval. This estimation procedure, which is rather complicated, has been used relatively rarely.

Source: Jenkins (2003: chapter 8).

With grouped duration data one can either assume that the durations are intrinsically discrete and treat them accordingly. Alternatively one may attempt to relate the model to an underlying process in continuous time. If one chooses the first option, one can use a "trick" demonstrated in Allison (1982). He shows that by assuming independent exits and a particular form of destination-specific hazards, the likelihood has the same form as the likelihood for a standard multinomial logit model. If one chooses the second option, things get more complicated as the shape of the continuous time hazard rate within each interval cannot be identified from the available grouped data. On top of the assumption of independent exits, the construction of the sample likelihood requires then assumptions about the shape of the continuous time hazard rate. Alternative assumptions lead to different econometric models. Jenkins (2003:chapter 8) has a detailed overview (with derivations) of some of the various estimation strategies that have been attempted in the literature. Table 4.1 summarizes his discussions.

The approach taken here is to treat the data as if they were discrete, thereby eschewing any attempt to relate the model to an underlying process in continuous time, c.f. the estimation strategy in Allison (1982). The way to proceed is to start off by assuming that the exits are independent of each other. The assumption of independent exits implies that the overall likelihood contribution for a firm can be written in the following way:

$$L = (L^{E1})^{\delta^{E1}} (L^{E2})^{\delta^{E2}} (L^{E3})^{\delta^{E3}} (L^{active})^{1-\delta_{E1}-\delta_{E2}-\delta_{E3}}$$

$$= \left[\frac{h_{E1}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)}\right]^{\delta_{E1}} \left[\frac{h_{E2}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)}\right]^{\delta_{E2}}$$

$$\times\left[\frac{h_{E3}(t)}{1-h_{E1}(t)-h_{E2}(t)-h_{E3}(t)}\right]^{\delta_{E3}} \times \prod_{a=1}^{t} [1-h_{E1}(a)-h_{E2}(a)-h_{E3}(a)]$$

where *a* denotes the day the firm gets incorporated, and where δ_{E1} = 1 when the specific firm exits because of E1, δ_{E2} = 1 when the specific firm exits because of E2, and δ_{E3} = 1 when the specific firm exits because of E3. When the firm does not exit for E1, E2 or E3 reasons, it is active (and gets censored in 2001).

Even though independent exits are assumed, the discrete-time likelihood cannot be factored into separate components for each of the three events. Maximum likelihood estimation must be done simultaneously for all kinds of events. It turns out, that by assuming that the destination-specific hazards have a particular form, there is still a straightforward way of estimating this independent competing-risks model, as the likelihood function can then be rewritten to have the same form as a standard multinomial logit model. This is shown in Allison (1982). Following Allison (1982) it is assumed that the destination-specific hazards are of the form

$$h_{E1}(t) = \frac{\exp(\theta_{E1} + \beta'_{E1} X_t)}{1+\exp(\theta_{E1} + \beta'_{E1} X_t) + \exp(\theta_{E2} + \beta'_{E2} X_t) + \exp(\theta_{E3} + \beta'_{E3} X_t)}$$

$$h_{E2}(t) = \frac{\exp(\theta_{E2} + \beta'_{E2} X_t)}{1+\exp(\theta_{E1} + \beta'_{E1} X_t) + \exp(\theta_{E2} + \beta'_{E2} X_t) + \exp(\theta_{E3} + \beta'_{E3} X_t)}$$

$$h_{E3}(t) = \frac{\exp(\theta_{E3} + \beta'_{E3} X_t)}{1+\exp(\theta_{E1} + \beta'_{E1} X_t) + \exp(\theta_{E2} + \beta'_{E2} X_t) + \exp(\theta_{E3} + \beta'_{E3} X_t)}$$

where θ characterizes the baseline-hazard function. The baseline-hazard function can be specified parametriclly and non-parametrically. A non-parametric specification is preferred as any inconsistency caused by misspecification is then avoided. In the estimations, dummies for each age (up to a dummy with companies that are 30 years old or older) are included (the reference category is firms that are 1 year old). The specification could be denoted $D_d'\gamma$ where D_d would then denote the dummies and γ would denote the coefficients to be estimated.

X_t characterizes the covariates (except of age which enters as described above), that are listed in section 2 in tables 2.1, 2.2 and 2.3. Note the difference between the age of the specific companies and the controls for the macroeconomic environment.

β are the parameters of the time-varying covariates.

Inserting the destination-specific hazards gives the following expression for $h_{active}(t)$:

$$h_{active}(t) = 1 - h_{E1}(t) - h_{E2}(t) - h_{E3}(t)$$

$$= \frac{1}{1 + \exp(\theta_{E1} + \beta'_{E1} X_t) + \exp(\theta_{E2} + \beta'_{E2} X_t) + \exp(\theta_{E3} + \beta'_{E3} X_t)}$$

In the above specifications, h_{E1} depends on E2 and E3, h_{E2} depends on E1 and E3, and h_{E3} depends on E1 and E2, and therefore, in the actual estimations, the competing risks (the three hazards) are not independent. Nonetheless Jenkins (2003) calls the model an independent competing-risks model, as the correlations between unobservable factors affecting each destination-specific hazard are assumed away. The expressions for the destination-specific hazards are inserted in the overall likelihood contribution for a specific firm, and the same form as the likelihood for a standard multinomial logit model is obtained. As before, δ_{E1} = 1 when the specific firm exits because of E1, δ_{E2} = 1 when the specific firm exits because of E2, and δ_{E3} = 1 when the specific firm exits because of E3. The likelihood is then:

$$L=\left[\frac{\exp(\theta_{E1}+\beta'_{E1}X_t)}{1+\exp(\theta_{E1}+\beta'_{E1}X_t)+\exp(\theta_{E2}+\beta'_{E2}X_t)+\exp(\theta_{E3}+\beta'_{E3}X_t)}\right]^{\delta_{E1}}$$

$$\times\left[\frac{\exp(\theta_{E2}+\beta'_{E2}X_t)}{1+\exp(\theta_{E1}+\beta'_{E1}X_t)+\exp(\theta_{E2}+\beta'_{E2}X_t)+\exp(\theta_{E3}+\beta'_{E3}X_t)}\right]^{\delta_{E2}}$$

$$\times\left[\frac{\exp(\theta_{E3}+\beta'_{E3}X_t)}{1+\exp(\theta_{E1}+\beta'_{E1}X_t)+\exp(\theta_{E2}+\beta'_{E2}X_t)+\exp(\theta_{E3}+\beta'_{E3}X_t)}\right]^{\delta_{E3}}$$

$$\times\left[\frac{1}{1+\exp(\theta_{E1}+\beta'_{E1}X_t)+\exp(\theta_{E2}+\beta'_{E2}X_t)+\exp(\theta_{E3}+\beta'_{E3}X_t)}\right]^{1-\delta_{E1}-\delta_{E2}-\delta_{E3}}$$

$$\times\prod_{a=1}^{t-1}\frac{1}{1+\exp(\theta_{E1}+\beta'_{E1}X_a)+\exp(\theta_{E2}+\beta'_{E2}X_a)+\exp(\theta_{E3}+\beta'_{E3}X_a)}$$

Left truncation and right censoring is handled as in Allison (1997:227) and Jenkins (1995). Note that it is chosen to set the censored firms (the active firms) as the reference category.[11]

The coefficients reported in the appendix on figures and tables (section 8) may be interpreted just like coefficients in a binary logit model, c.f. Allison (2001:119). In binary logit analysis, if a variable x has a positive coefficient, one knows that every increase in x results in an increase in the probability of the designated outcome. When estimating the multinomial logit model it is tempting to describe the effects of the explanatory variables on the various hazards as the increase the respective variables produce on the probability of being in the E1, E2 and E3 category respectively. But as it is seen from the equations below (which are the ones that are actually estimated in the multinomial logit model) the effects must be interpreted as contrasts between pairs of categories.

$$\log\left(\frac{h_{E1}}{h_{active}}\right)=\theta_{E1}+\beta'_{E1}X_t$$

[11] When estimating a multinomial logit model, one must be aware of the identification issue. The *m* alternative choices (E1, E2, E3 and being censored) cannot be identified, as there is more than one set of estimates that would lead to the same probabilities of the outcomes observed. The way to deal with this problem is to set one of the sets of parameters equal to zero.

$$\log(\frac{h_{E2}}{h_{active}}) = \theta_{E2} + \beta'_{E2} X_t$$

$$\log(\frac{h_{E3}}{h_{active}}) = \theta_{E3} + \beta'_{E3} X_t$$

This means for example that in the competing-risks model the odds that a private limited liability company will enter financial distress rather than staying active are about exp(0.4174)=1.52 the odds for public limited liability companies. For the solvency ratio which has exp(-2.5103)=0.08, each 1-level increase in the variable multiplies the odds of moving into financial distress versus staying active by about 0.08.

The fact that the multinomial logit coefficients always must be interpreted as effects on contrasts between pairs of categories and not on the probability of being in a particular category is often referred to as “relative risk” or as the assumption of “independence of irrelevant alternatives” (IIA). The assumption implies that adding another alternative (or changing characteristics of another alternative) does not affect the relative odds between two of the alternatives. The IIA assumption is implausible for applications with similar alternatives. Wooldridge (2002) and Train (2003) consider the well-known red bus/blue bus example, where adding a blue bus to the choice set of taking the train or the red bus, changes the probability of taking the train, even though the probability of taking the train should be the same as before and not depend on the new choice – a blue bus. The simple example is this: first there is the choice of taking a red bus or the train instead of a car. Say, that the estimated probability of taking the red bus/car is one half and so is the train/car probability. Now, a new alternative is added, e.g. a blue bus. The IIA property implies that the relative odds between taking the red bus and the train are the same as before. The new probabilities will then be the following: the probability of taking the blue bus/car is one third, the red bus/car is one third and also that the probability of taking the train/car is one third (note that the relative odds between taking the red bus and the train would not be changed). But this makes no sense. It should not matter whether the bus is red or blue. It would be more realistic if the probability of taking the train/car was still one half, and then the probability of taking the red bus/car and blue bus/car were both, one fourth.

The potential problem sketched in the example, is not something to worry about in the problem of interest in this paper. Here the alternatives E1, E2 and E3 are not similar, and so, a priori, there should be no reason to believe that one category could be excluded. Nonetheless, to be sure, it is chosen to test the assumption of independence of irrelevant alternatives. The assumption can be tested with

Hausman's specification test, c.f. Greene (2003:725)., who writes that "if a subset of the choice set truly is irrelevant, omitting it from the model altogether will not change parameter estimates systematically. Exclusion of these choices will be inefficient but will not lead to inconsistency. But if the remaining odds are not truly independent of these alternatives, then the parameter estimates obtained when these choices are included will be inconsistent". The test statistic is $\chi^2 = (\hat{\beta}_s - \hat{\beta}_f)'[\hat{V}_s - \hat{V}_f]^{-1}(\hat{\beta}_s - \hat{\beta}_f)$, where s indicates the estimators based on the restricted subset and f indicates the estimator based on the full set of choices, and $\hat{V}$ are the estimates of the asymptotic covariance matrices. The statistic has a limiting chi-squared distribution with K degrees of freedom (i.e. the number of parameters to be tested). When testing for the IIA assumption, the conclusion is that when E2 and E3 firms are left out, respectively, the difference in coefficients is not systematic. This implies that the IIA specification holds and therefore that the model is correctly specified.

To account for potential unobserved characteristics in the multinomial logit model, unobserved heterogeneity can be introduced. Unobserved heterogeneity is not introduced here as the problem in this paper is a forecasting problem (as opposed to a problem of econometric measurement), c.f. the discussion in Arellano (2003:11). Instead of including unobserved heterogeneity, proxy variables are suggested, as was discussed in section 2.2. Jenkins (2003:102) notes that empirical work suggests, that the effects of unobserved heterogeneity are mitigated, and thence estimates more robust, if a flexible baseline-hazard specification is used (as it is in this case), and that the topic of unobserved heterogeneity underscores the importance of getting good data, including a wide range of explanatory variables that summarize well the differences between, in this case, the firms. This means that if one has a wide range of explanatory variables, it is less of a problem not to control for unobserved heterogeneity.[12]

---

[12] It is not straightforward to introduce unobserved heterogeneity in this model. When incorporating unobserved heterogeneity in problems like this one, the usual assumption is that the unobserved individual effect is destination state specific, time-constant, and independent of the observed characteristics. Malchow-Møller and Svarer (2002) derive the likelihood and discuss how the random effects multinomial logit model can be estimated using SAS. They go through the procedure and underline that it works well for small models and/or samples, and that it might work surprisingly slowly for larger models/samples.They mention that using a data set with a total of 6,927 observed choices it took a PC of type Pentium III with 733 MHz and 128 RAM more than 30 hours to estimate a model with more than 3 alternative choices and more than 60 days to estimate a model with 4 alternative choices. Here there are 4 alternative choices and more than 150,000 observations!

4.2. Estimation Results

If one assumes that the exits are independent, and that the special kind of destination-specific hazard rates, which were discussed in section 4.1 are the true hazards then a competing-risks model can be estimated as a standard multinomial logit model. The result from the estimation of the competing-risks model with these assumptions is presented in this section. It should be emphasized that if the specification of the model does not correspond to the true model, then the parameter estimates will not be consistent. Note that the assumption of independent exits is necessary for identification, and that it therefore cannot be tested.

The global tests (wald tests, not reported) for the effect of each variable on the outcome variable, controlling for the other variables in the model, shows that none of the core variables and only two proxies, namely the concentration index and the dummy for being owned by a fund, have no effect of the outcome variable. When testing (likelihood ratio test, not reported) whether the E1 coefficients are identical to the E2 and E3 coefficients, respectively, the null hypothesis are rejected, implying that at least one pair of coefficients differs across the equations. This means that at least one pair of coefficients differs between E1 and E2 coefficients, and that at least one pair of coefficients differs between E1 and E3 coefficients.

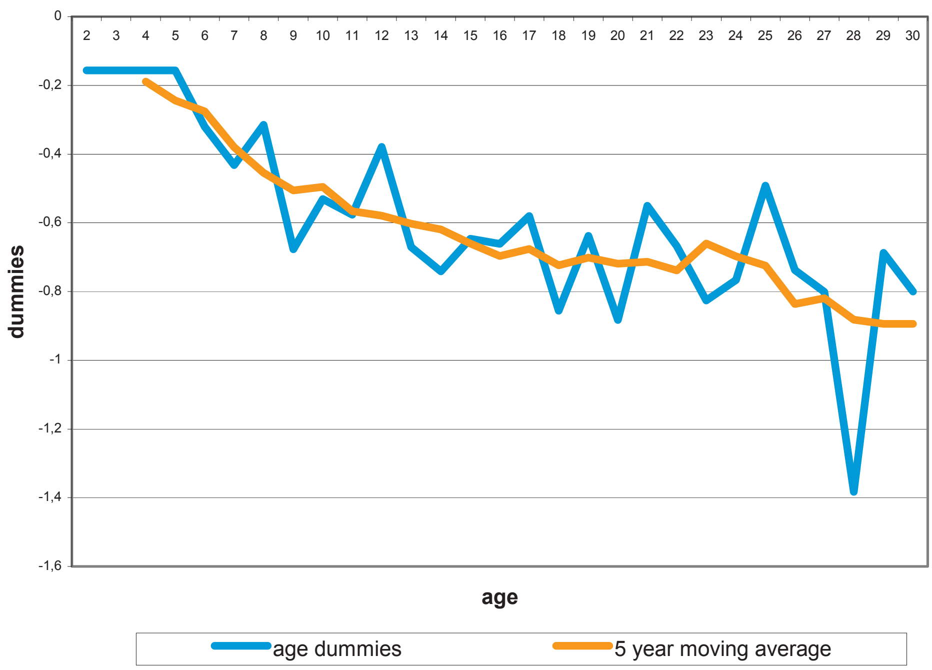
The results on the E1 hazard (relative to staying active) from the estimation of the coefficients in the parametric competing-risks model are seen from tables 4.2.a and 4.2.b. The tables show that in all cases where an effect of a parameter was hypothesized, the parameter estimates have the expected sign when the competing-risks model is estimated, except in the case of the effect of duration dependence, c.f. figure 4.2.a. The duration dependence was hypothesized to be bell-shaped, but the effect is almost linear until the firms reach the age of 15, and from the age of 15 the effect is approximately constant.

Table 4.2.a: Core variables (E1 parameter estimates)

Variables	Estimated effect	Expected effect
Firm age (dummies)	See the text and figure 4.2.a	Bell-shaped effect
Short term debt to total assets	Positive*	Positive
Return on net assets	Negative*	Negative
Solvency ratio	Negative*	Negative
Firm size	Negative*	Negative

Note: The variables are significant at the 1 per cent level (indicated by *). There is controlled for the macroeconomic environment and for the various sectors. In the estimations, farming and forestry is included in the same sector affiliation category, as the data were too sparse otherwise. The same holds for mining, energy and construction. The IT dummy was positive and significant at the 1 per cent level. The primary bank categories have been altered: as the data was too sparse otherwise, firms that register a bank in category 3 or 4 are now in the same group. Only firms that reported a primary bank in category 2 had a significantly larger probability of entering financially distress. The other bank dummies were not significantly different from zero. The dummy for publicly traded companies is not included in the estimations as no publicly traded firms enter financial distress in the period.

Figure 4.2.a: Firms in financial distress: duration dependence



Note: The figure sketches the age dummies (reference dummy is firms that are equal to 1 year old). The last dummy is called 30 years old or older. All the reported dummies are significant at the 5 per cent. level. Most of them are also significant at the 1 per cent level.

For some of the variables a particular sign was not expected. Instead it was left to the estimations to show whether or not the variable was significant, and what sign the parameter has. This was the case for the dummy variables ultimate parent company, owned by a fund, owned by the public sector, local authority group 2, 3, 4 and 5, and for the control variable primary bank group 1, group 2, and group 3 & 4 (the estimates on the control variables are not reported).

The estimation of the competing-risks model shows that the sign of the parameter of the dummy variable “ultimate parent company” is positive, and so the odds, that an ultimate parent company will enter financial distress relative to staying active compared to companies that are not ultimate parent companies, are higher. The reason for this may be that the ultimate parent companies have a tendency to let the not-so-well-performing subsidiaries drain resources from the group as a whole, including the ultimate parent company, e.g. because of reputational risk.

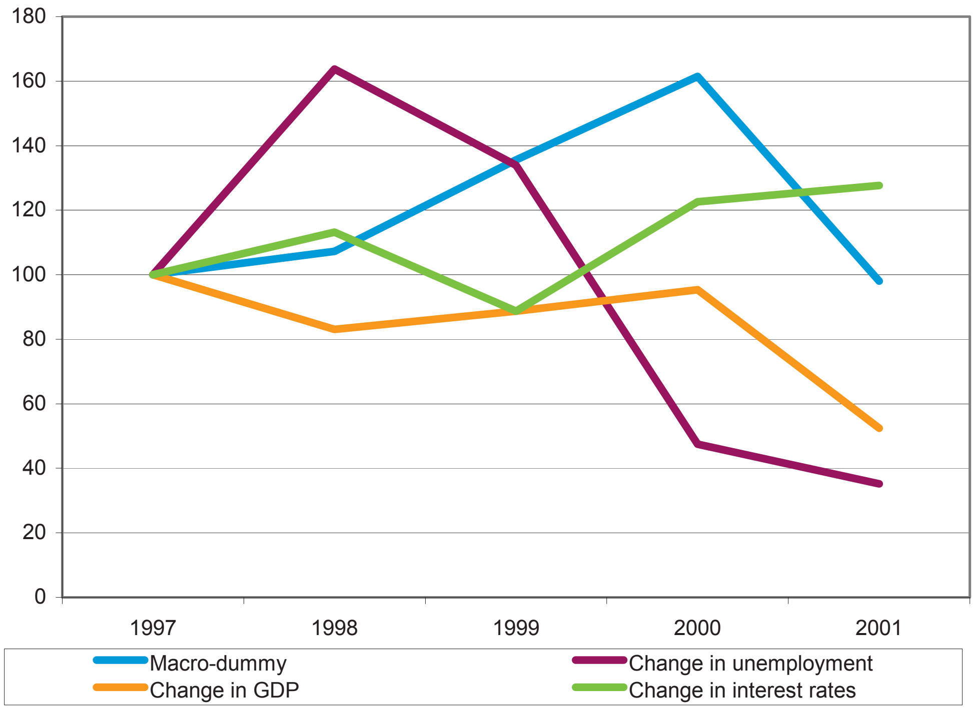
The effects of fund ownership and public sector ownership are not significant. This is in line with the discussion in section 2, where the conclusion was that it was left to the estimations to show whether there is a significant effect of these ownerships variables at all. The effect of the location of a firm is not clear a priori. The estimations of the competing-risks model show that companies in all local authority groups, except of local authority group 4, have smaller odds of moving into financial distress compared to the companies situated in the Copenhagen area (local authority group 1). The odds that firms that register a group 2 bank, as their primary bank enter financial distress are higher than the odds for firms that do not register a bank connection. There is no difference between not registering a bank at all and registering a group1 or group3&4 banks.

Table 4.2.b: Proxies

Variables	Estimated effect	Expected effect
Owned by the public (dummy)	Not sign.	?
Owned by a fund (dummy)	Not sign.	?
Diversification 2 sectors (related business) (dummy)	Negative*	Negative
Diversification 3–9 sectors (related business) (dummy)	Negative*	Negative
Diversification 2 sectors (unrelated business) (dummy)	Negative*	Negative
Diversification 3–9 sectors (unrelated business) (dummy)	Negative*	Negative
Local authority group 1 (reference dummy)		
Local authority group 2 (dummy)	Negative*	?
Local authority group 3 (dummy)	Negative*	?
Local authority group 4 (dummy)	Not sign.	?
Local authority group 5 (dummy)	Negative*	?
Concentration	Not sign.	?
Critical comments from the auditors (dummy)	Positive*	Positive
Ultimate parent companies (dummy)	Positive**	?
Wholly-owned subsidiaries (dummy)	Negative*	Negative
Private limited liability company		
Public limited liability company (reference dummy)	Positive*	Positive

Note: * indicates that the variable is significant at the 1 per cent level. ** indicates that the variable is significant at the 5 per cent level. See also the note to table 4.2.a.

Figure 4.2.b: Macroeconomic dummies and macroeconomic variables



Note: The macroeconomic dummies as well as the change in unemployment, the change in GDP and the change in interest rates are normalized with 1997 as the base year (which is set equal to 100).

The dummy variables that control for the sector affiliation in the competing-risks model show that compared to being a manufacturing firm (reference dummy) the odds that the firms that belong to one of the following sectors move into financial distress are smaller: trade and hotel, transport, business service, public service activities, and organizations, etc. On the contrary, the odds that the firms that have the sector affiliation "unknown" move into financial distress are higher. Firms with the other sector affiliations are not significantly different from firms in the manufacturing sector. All firms that are in the self-constructed IT- and telecommunication category have higher odds of moving into financial distress than all other firms. The macroeconomic dummies as well as macroeconomic variables are sketched in figure 4.2.b.

4.3. Goodness-of-fit and Robustness

The goodness-of-fit of the model can be judged from table 4.3.a. The table tabulates the actual events versus the average predicted probability of the respective events. No cut-off level is chosen. The predicted event is the one with the highest probability. The important thing to notice is that the diagonal numbers are the largest. This implies that the model can distinguish between the various exits.

Table 4.3.b gives an idea about the robustness over time. As expected, for all estimated periods, the diagonal numbers are the largest. It is notable that the model estimated using the whole period 1995 – 2001 seems to be the “worst” model. The explanation can be the increase in the number of firms in financial distress (i.e. increased volatility) in the period after 1998, c.f. figure 3.2.

Table 4.3.b: Goodness-of-fit using five different samples

Goodness-of-fit		Actual				
		E1	E2	E3	Active	All
Average predicted probability	E1	**0.15**	0.05	0.02	0.01	0.01
		0.14	0.05	0.02	0.01	0.01
		0.12	0.04	0.02	0.01	0.01
		0.12	0.04	0.02	0.01	0.02
		0.10	0.03	0.02	0.01	0.02
	E2	0.01	**0.11**	0.01	0.00	0.00
		0.01	**0.11**	0.01	0.00	0.00
		0.01	**0.10**	0.01	0.00	0.00
		0.01	**0.09**	0.01	0.00	0.00
		0.01	**0.08**	0.01	0.00	0.01
	E3	0.01	0.03	**0.08**	0.00	0.01
		0.01	0.03	**0.06**	0.01	0.01
		0.01	0.03	**0.05**	0.01	0.01
		0.01	0.02	**0.04**	0.01	0.01
		0.01	0.02	**0.04**	0.01	0.01
	Active	0.82	0.81	0.89	**0.98**	0.98
		0.84	0.81	0.90	**0.98**	0.98
		0.86	0.83	0.92	**0.98**	0.98
		0.87	0.84	0.93	**0.98**	0.97
		0.88	0.87	0.94	**0.97**	0.97

Note: This table pictures the average predicted probability estimated using six different samples. The samples are sketched in the following order: Sample 1995 – 1997 (first line), sample 1995 – 1998 (second line), sample 1995 – 1999 (third line), sample 1995 – 2000 (fourth line) and sample 1995 – 2001 (fifth line). For further details on how to read the table see the note to table 4.3.a.

Table 4.3.a: Goodness-of-fit: Competing-risks model

Goodness-of-fit		Actual				
		E1	E2	E3	Active	All
Average predicted probability	E1	**0.1019**	0.0340	0.0178	0.0139	0.0154
	E2	0.0108	**0.0800**	0.0098	0.0046	0.0051
	E3	0.0074	0.0199	**0.0360**	0.0070	0.0073
	Active	0.8798	0.8661	0.9384	**0.9736**	0.9723

Note: The table should be read in the following way: out of the firms that end up in financial distress (see the "actual E1" column), 10.19 per cent were predicted to end up in financial distress, 1.08 were predicted to end up as voluntarily liquidated, 0.74 were predicted to merge with other companies or the like, and 88 pct. were predicted to be active. Note that all columns sum to 1. Look now at the predicted row: 10.19 per cent of the firms that end up as financially distressed firms were predicted to enter financial distress, 3.4 per cent of the firms that end up as voluntarily liquidated were predicted to enter financial distress, 1.78 per cent of the firms that end up as mergers and the like were predicted to enter financial distress, and 1.39 per cent of the firms that end up as active were predicted to enter financial distress.

4.4. A Comparison of different Model Specifications

In this study a parametric competing-risks model was set up, and the exit to E1, E2 and E3 was modelled. Other studies pool the various exit types, e.g. Bunn (2003) who discusses the company-accounts-based modelling of business failure in Bank of England. In the paper, which uses the probit model to assess the probability that a company fails, a firm is defined to fail in a particular year if its company status is in receivership, liquidation or dissolved (this means that the definition includes voluntary liquidation and dissolution where there may be no risk of default).

As papers are still published where the various exit types are pooled, it is relevant to compare the results in the parametric competing-risks specification and a pooled logit regression. In the pooled logit regression E1, E2 and E3 exits are pooled into one group (called E_pooled), and the event E_pooled is modeled.

In the papers by Ohlson (1980) and Shumway (2001), bankrupt firms are modeled. In order to compare the set up in their models with the competing-risks model, it is chosen to model the E1 event, and to treat E2 and E3 events in the same way as active firms are treated, i.e. to right censor all other firms but E1 firms. The model is denoted the E1event model.

Below the competing-risks model is compared to the pooled logit model and the E1event model. Firstly the sign of the parameter estimates in the various specifications is discussed. Secondly the proportion of correct predictions is discussed.

4.4.1. The sign of the Parameter Estimates

The conclusion from a comparison of the results from the competing-risks model and the pooled logit model is that it is important to distinguish between exit types. The sign of the parameter estimates of the E1 competing-risks hazard differ from the sign of the parameter estimates of the pooled logit model (c.f. the appendix on figures and tables in section 8). This result is not surprising as the likelihood ratio test in section 4.2 shows, that when the E1 coefficients are compared to the E2 and E3 coefficients, respectively, at least one pair of coefficients differs across the equations. The result is in line with Harhoff et al. (1998:470) who consider two (not three) exit modes and conclude that "Introducing the distinction between insolvency and voluntary liquidation clearly calls these results [the pooled logit results] into doubt and reveals that pooling exit types is a major source of misspecification."

In the pooled logit regression, the duration dependence is different from the duration dependence, which was found when estimating the competing-risks model (in the pooled logit model only three age dummies are significant: 1) dummy: age between 2 and 18, 2) dummy: 19 years, and 3) dummy: 20 years and above). Also, the sign on the wholly-owned subsidiary dummy changes from negative to positive, and the ultimate parent dummy is not significant in the pooled logit model, whereas it was significant and positive in the competing-risks model. Furthermore the legal status dummy (which takes the value one for private limited liability companies) is not significant in the pooled logit estimations, whereas it was significant and positive in the competing-risks model. Other variables that have different signs in the two different specifications are owned by the public (is not significant in the competing-risks model and is positive in the pooled logit model), belonging to local authority group 4 (which is not significant in the competing risks model and is negative in the pooled logit model), and the primary bank dummy for group 2 banks (which is significant and positive in the competing-risks model and insignificant in the pooled logit regression).

The sector affiliation dummies also have different signs in the two specifications. In the pooled logit regression only the dummy on the sectors unknown and not stated were significant (and positive). In the competing-risks model some dummies were significant and negative (trade and hotel, transport, business service, public service activities, and organizations, etc.) and only the dummy on the sector affiliation "unknown" was significant and positive.

When comparing the sign of the parameter estimates of the financial distress hazard in the competing-risks model and the parameter estimates in the E1event model, it is seen that all parameter estimates have the same sign, but that the magnitude differs, c.f. the appendix on figures and tables (section 8). One reason for this result could be that in the sample there are not many E2 and E3 firms compared to the number of active firms. The variables with the largest differences in magnitude are the return on net assets and the solvency ratio. It is important to note, that when one estimates the E1event model, one does not obtain estimates of the E2 and the E3 hazard.

4.4.2. The Proportion of Correct Predictions

A measure of how well the model with the specifications for the hazard function fits the data is the proportion of correct predictions. The naïve predictor uses a cut-off value of 0.5, which means that firms with a predicted probability above 0.5 are classified as financially distressed firms, whereas firms with a predicted probability below or equal to 0.5 are classified as active firms. Even though 0.5 is the usual choice, it may not be a proper value to use for the threshold in all cases. In this case, where the sample is unbalanced, i.e. where there are few events (=financially distressed firms) compared to non-events, the 0.5 prediction rule may never predict a financially distressed firm as being financially distressed. The proportion of financially distressed firms used in the estimations to all other firms (E2 and E3 firms, as well as active or censored firms) is 0.016, that is, for every 1,000 active or censored firm there is in the sample, there are 16 financially distressed firms. In this setting, it may require an extreme configuration of regressors even to produce a predicted probability of 0.03, not to say 0.5 (Greene (2003:685)), and so it seems natural to reduce the cut-off level from 0.5, in order to predict financially distressed firms more often. It will, of course, increase the number of times that active firms and E2 and E3 firms are incorrectly classified as being financially distressed firms. In fact, this is the trade-off that one has to make. The question to be answered is how bad it is to incorrectly classify a firm that does not exit because of financial distress as a financially distressed firm compared to not classifying a financially distressed firm as financially distressed. When the cut-off level is changed, it will always reduce the probability of one type of error while increasing the probability of the other.

As the proportion of financially distressed firms used in the estimations of the competing-risks model to all other firms is 0.016, as a start, a cut-off level of 0.016 is used. With this cut-off level the competing-risks model correctly classifies 78 per cent of the financially distressed firms as financially distressed, c.f. table 4.4.2.a, whereas the proportion of correctly called non-events is 80 per cent. Table 4.4.2.b reports the proportion of predicted events split up on model prediction and actual exit. The tables shows that 31 per cent of E2 firms and 22 per cent of E3 firms are predicted to be in financial distress, but end up as E2 or E3. Based on table 4.4.2.b., it seems reasonable to group E2, E3 and active firms.

Table 4.4.2.a: Competing-risks model

	Model prediction: Financial distress (event)	Model prediction: Non-event (corresponding to either E2, E3 or an active firm)
Financial Distress	Correct call of event: 78 pct. (2,024 out of 2,586)	Type 1 error: Missing prediction: 22 pct. (562 out of 2,586)
E2, E3 or active firm	Type 2 error: Wrong signal: 20 pct. (33,039 out of 165,764)	Correct call of non-event: 80 pct. (132,725 out of 165,764)

Note: The number of firms is not exactly the same as the number of firms reported in section 3.2, e.g. 2,586 firms in financial distress are included in the estimations instead of the original sample of 2,617 firms in financial distress. As mentioned in section 3.2 this is due to a small number of missing observations in the final estimations. As a cut-off level 0.016 is used.

Table 4.4.2.b: Competing-risks model (E1, E2, E3 and active firms split up on actual exit)

	Model prediction: Financial distress (event)	Model prediction: Non-event (corresponding to either E2, E3 or an active firm)
E1 (financial distress)	Correct call of event: 78 per cent (2,024 out of 2,586)	Type 1 error: Missing prediction: 22 per cent (562 out of 2,586)
E2 (voluntary liquidation)	Type 2 error: Wrong signal: 31 per cent (266 out of 856)	Correct call of non-event: 69 per cent (590 out of 856)
E3 (mergers/acquisitions etc.)	Type 2 error: Wrong signal: 22 per cent (272 out of 1,224)	Correct call of non-event: 78 per cent (952 out of 1,224)
Active	Type 2 error: Wrong signal: 20 per cent (32,501 out of 163,684)	Correct call of non-event: 80 per cent (131,183 out of 163,684)

Note: See the note to table 4.4.2.a on the number of firms in the final estimations. As a cut-off level 0.016 is used.

In figure 4.4.2.a the predicted value of the firms that are predicted to end up in financial distress (i.e. have a predicted value above 0.016) are sketched. The predictions are split up on the actual exit. This means that what can be read from the table is *the actual exit of the firms that are predicted to end up in financial distress*. This is important. According to the model the firms are predicted to end up as financially distressed. But because the actual exit is known, the firms that are predicted as financially distressed can be grouped according to the actual exit. The predictions are split up on chosen percentiles (1, 10, 25, 50, 75, 90 and 99 per cent). Take one of them, e.g. the 90th percentile. As is the expectation, the figure shows that at the 90th percentile, firms that actually end up in financial distress have a higher predicted value than firms that actually end up as E2, E3 or active firms. The more interesting thing to note is that firms that end up exiting for E2 or E3 reasons (but are predicted to enter financial distress) have higher predicted values than the firms that end up as active firms (but are predicted to enter financial distress).

To conclude, figure 4.4.2.a shows that the firms that actually end up as E2 and E3 firms have predicted values that are higher than the predicted values of the firms that end up as active firms, or in other words, compared to the active firms that are predicted to end up in financial distress (the so called worst faring active firms), the E2 and E3 firms that are predicted to end up in financial distress (the so called worst faring E2 and E3 firms) have predicted values that are closer to the predictions that the E1 firms generate.

Figure 4.4.2.b shows the predicted value of the firms that are predicted to end up as active (i.e. have a predicted value below 0.016). The predictions are split up on the actual exit. Again this means that what can be read from the figure is *the actual exit of the firms that are predicted to end up as active*. The figure shows that the firms that actually end up as E2 or E3 firms have predicted values that are smaller than the predicted values of the firms that actually end up as active firms. Compared to E2 and E3 firms, active firms have predicted values that are closer to the predictions that the E1 firms generate (This is the case for all the chose percentiles in figure 4.4.2.b, except for the 99 percentile for the E3 firms).

The overall conclusion from figures 4.4.2.a and 4.4.2.b is that it seems as if the worst faring E2 and E3 firms (the ones that are predicted to enter financial distress, c.f. figures 4.4.2.a) are “weaker” than the worst faring active firms, and that the best faring E2 and E3 firms (the ones that are predicted to be active, c.f. figure 4.4.2.b) are “stronger” than the best faring active firms.

If the cost of not predicting more events is high, one can reduce the cut-off level used in table 4.4.2.a in order to predict more events. This will be at the cost of an increased number of type 2 errors (wrong signal). A reduction in the cut-off level to 0.008 increases the proportion of correctly called events to 91 per cent and the proportion of type 2 errors to 42 per cent. If one judges, that the cost of making too many type 2 errors is high, one can increase the cut-off level in order to make less type 2 errors. This will be at the cost of less correctly called events. If one increases the cut-off level to 0.04, the proportion of type 2 errors reduces to 6 per cent and the proportion of correctly called events is reduced to 50 per cent.

Figure 4.4.2.a: Firms predicted as financially distressed in the competing-risks model: The distribution of predictions split up on E1, E2, E3 and active firms

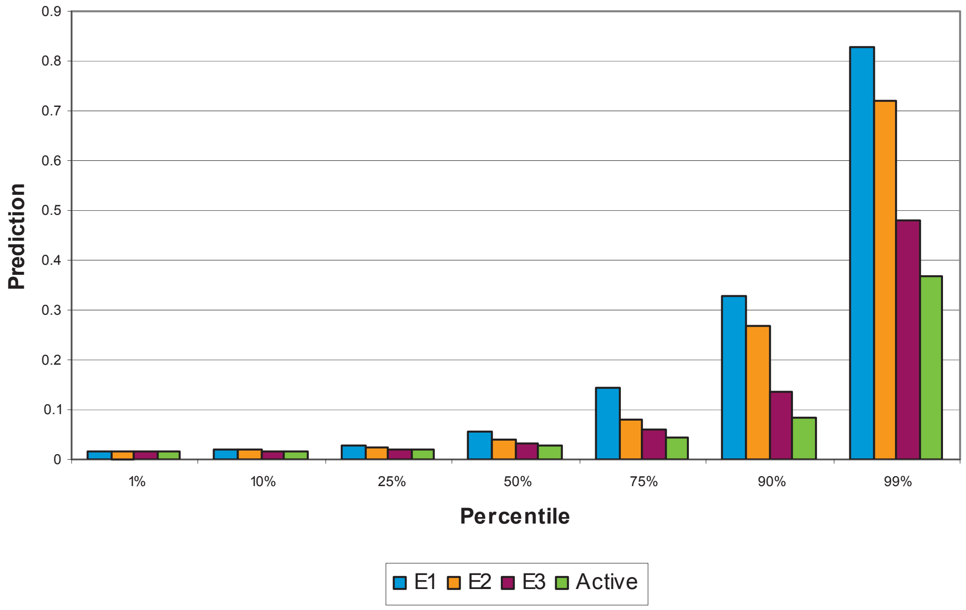


Figure 4.4.2.b: Firms predicted as non-events in the competing-risks model (E2, E3 and active): The distribution of predictions split up on E1, E2, E3 and active firms

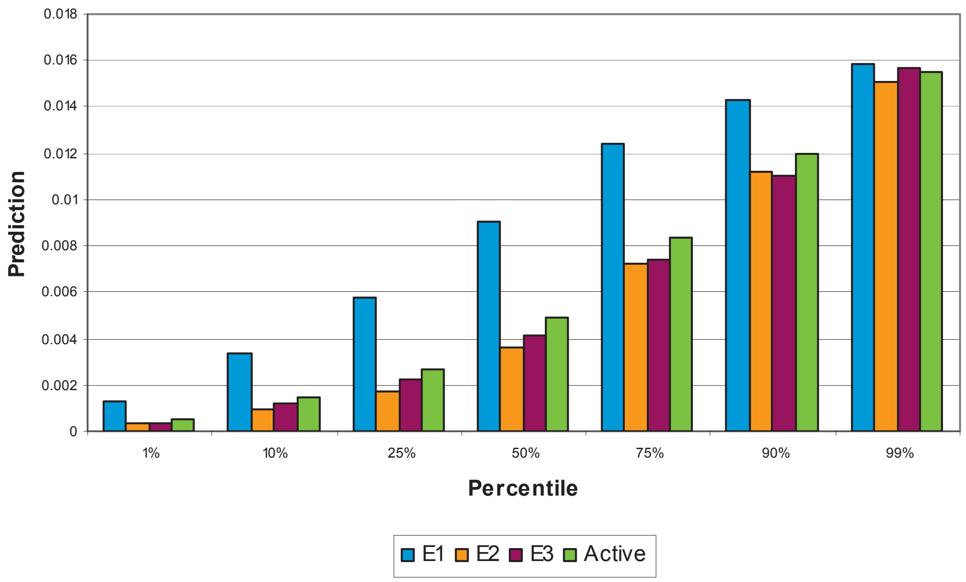


Table 4.4.2.c: Pooled logit regression

	Model prediction: Event (=an event)	Model prediction: Non-event (=active firms)
Event (=an exit)	Correct call of event: 65 pct. (3,036 out of 4,666)	Type 1 error: Missing prediction: 35 pct. (1,630 out of 4,666)
Non-event (=active firms)	Type 2 error: Wrong signal: 25 pct. (41,445 out of 163,684)	Correct call of non-event: 75 pct. (122,239 out of 163,684)

Note: The proportion of events to non-events is 0.029 (corresponding to 4,666 out of 163,684), and therefore 0.029 is used as the cut-off level. See also the note attached to table 4.4.2.a. concerning missing data in the final estimations.

Table 4.4.2.d: E1event model

	Model prediction: Event (=an event)	Model prediction: Non-event (=active firms)
Event (=an exit)	Correct call of event: 78 pct. (2,026 out of 2,586)	Type 1 error: Missing prediction: 22 pct. (560 out of 2,586)
Non-event (=active firms)	Type 2 error: Wrong signal: 20 pct. (33,140 out of 165,764)	Correct call of non-event: 80 pct. (132,624 out of 165,764)

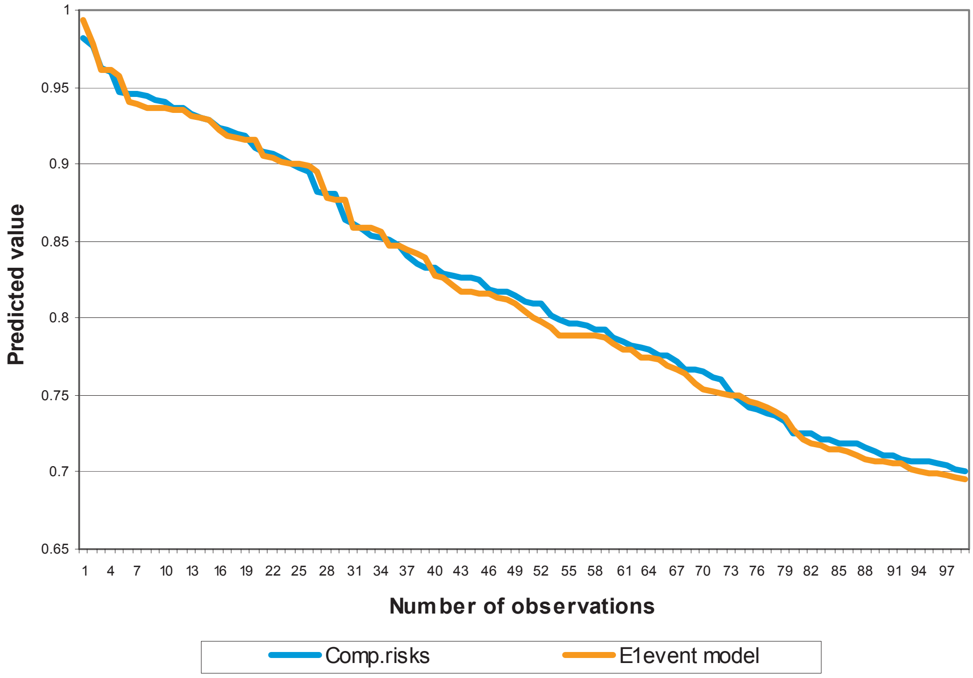
Note: The proportion of events to non-events is 0.016, and therefore 0.016 is used as the cut-off level. See also the note attached to table 4.4.2.a. concerning missing data in the final estimations.

Table 4.4.2.e: E1event model (E1, E2, E3 and active firms split up on actual exit)

	Model prediction: Financial distress (event)	Model prediction: Non-event (corresponding to either E2, E3 or an active firm)
E1 (financial distress)	Correct call of event: 78 per cent (2,026 out of 2,586)	Type 1 error: Missing prediction: 22 per cent (560 out of 2,586)
E2 (voluntary liquidation)	Type 2 error: Wrong signal: 33 per cent (286 out of 856)	Correct call of non-event: 67 per cent (570 out of 856)
E3 (mergers/acquisitions etc.)	Type 2 error: Wrong signal: 23 per cent (278 out of 1,224)	Correct call of non-event: 77 per cent (946 out of 1,224)
Active	Type 2 error: Wrong signal: 20 per cent (32,576 out of 163,684)	Correct call of non-event: 80 per cent (131,108 out of 163,684)

Note: See the note to table 4.4.2.a on the number of firms in the final estimations. As a cut-off level 0.016 is used.

Figure 4.4.2.c: The 100 observations with the highest prediction in the competing-risks model and the E1event model



The goodness-of-fit analysis of the competing-risks model (table 4.4.2.a) is compared to the goodness-of-fit of the pooled logit model and the E1event model, which is reported in tables 4.4.2.c and 4.4.2.d, respectively. In the pooled logit model 65 per cent of the events are predicted to be an event, whereas the E1event model predicts 78 pct. of the financially distressed firms to be financially distressed. The number of correctly called non-events is 75 per cent in the pooled logit model and 80 pct. in the E1event model. When comparing the proportion of correct predictions in the competing-risks model with the pooled logit model, the conclusion is that the competing-risks model fare best, and so the conclusion is that it is important to distinguish between exit types.

The overall result when comparing the proportion of correct predictions in the competing-risks model with the E1event model is not as clear. The models generate predictions that are very similar. Tables 4.4.2.a and 4.4.2.d show that the E1event model correctly classifies 132,624 non-events, whereas the competing-risks model correctly classifies 132,725. Tables 4.4.2.b and 4.4.2.e, which provide more details, show that the number of correctly classified E2 firms is 590 in the competing-risks specification and 579 in the E1event model, and that the number of correctly classified E3 firms is 952 in the competing-risks specification and 946 in the E1event model. Concerning the classification of events, the E1event model correctly classifies 2,026 as financially distressed, whereas the competing-risks model correctly classifies 2,024 as financially distressed.

In figure 4.4.2.c the 100 observations with the highest predicted value in the competing-risks model and the E1event model are sketched. The figure suggests that the predictions are alike in the two specifications, i.e. the same conclusion as above. The conclusion is not changed when looking at the whole distribution of the predictions in the two model specifications, c.f. the appendix on predictions (section 9).

5. Conclusion

The primary goal of this paper is to create a model that can predict the firms that end up in financial distress. As the firms in the non-financial sector can exit for various reasons (financial distress, voluntary liquidation and because they are merged or acquired etc.) the method of competing-risks models seems appropriate, and so a parametric competing-risks model was considered.

Following Allison (1982) independent exits and a special kind of destination-specific hazard rates were assumed, and a parametric competing-risks model was estimated. The parametric competing-risks model was estimated using a panel data set containing information on the whole population of Danish non-financial public limited liability companies and private limited liability companies that existed between 1995 and 2001, covering around 30,000 firms and more than 150,000 firm-year observations. After the application of certain criteria (e.g. exclusion of holding companies and financial institutions), the sample is representative. More than 20 explanatory variables are included in the estimations.

One of the important things to note is that in the estimations, proxies for several of the variables that are inherently unobservable were used. In an estimation problem like this, the use of proxies are preferred (over a specification that leaves some of the variation to be modelled by an unobserved heterogeneity term), c.f. the discussion in Arellano (2003:11). Proxies are important. Take the example of a potential endogenous variable: the solvency ratio, which is calculated as equity capital over total assets. As the firms choose themselves the level of equity capital (e.g. how much they pay to their shareholders), whether they pay more or less to their shareholders might very well be correlated with uncertainty. This is usually a problem as uncertainty is not included in the estimations. Here it is less of a problem, simply because the four diversification variables as well as the location dummies and the concentration index are used as proxies for the uncertainty that the firms are facing.

The effects of the various explanatory variables (core variables, proxies and controls) in the competing-risks model are discussed. All the variables, except age, have the expected signs. For some variables a specific sign is not expected. The sign is estimated, presented and discussed.

The sign of the competing-risks E1 parameter estimates (the firms that enter financial distress) is compared to the sign of the parameter estimates from the estimation of a pooled logit model. Several variables do not have the same sign. When the competing-risks E1 parameter estimates are compared to the parameter estimates in the E1event model, the differences are in the magnitude of the parameter estimates. The sign of the E1 parameters are the same in the two model specifications.

A comparison of the predictive ability of the competing-risks model and the pooled logit model shows that the competing-risks model fare best, and so the conclusion is that it is important to distinguish between exit types. A comparison of the predictive ability of the competing-risks model with the E1event model is not as clear. The models generate predictions that are very similar. Note that when estimating the E1event model one does not obtain estimates of the E2 and the E3 hazard. If they are of interest one should estimate the competing-risks model.

The overall conclusions are the following: If the sign of the coefficients to the explanatory variables are of interest, then one should estimate the competing-risks model. This means that the conclusion in the Harhoff et al. (1988) paper, which distinguishes between two and not three exit modes, is verified, as long as the coefficients to the explanatory variables are of interest. If one, on the other hand, focuses on the predictive ability of the models, the conclusion is, that the simple financial distress model (where the exit to financial distress is modelled treating all other firms as censored) performs just as well as the parametric competing-risks model, and therefore, that if prediction is the sole purpose of the model, it does not matter which of these two models is estimated. Of course, the real test of the models' predictive abilities will be the practical confrontation with data for recent periods as it emerges. The predictive performance of the pooled logit model (where all exits are pooled) is far worse than both of the other two models.

6. LITERATURE

D'Addio, A. and M. Rosholm, 2002. *Left-Censoring in Duration Data: Theory and Applications*. Working Paper No. 2002-5, Department of Economics, School of Economics and Management, University of Aarhus

Agarwal, R. and D. B. Audretsch, 2001. Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival. *The Journal of Industrial Economics*, vol. XLIX, no. 1, pp. 21-43

Allison, P. D., 1982. Discrete-time Methods for the Analysis of Event Histories. In Samuel Leinhardt (ed.), *Sociological Methodology 1982*, San Francisco: Jossey-Bass, pp. 61-98

Allison, P. D., 1997. *Survival Analysis using the SAS System: A Practical Guide.* Cary, NC, USA: SAS Institute

Allison, P. D., 2001. *Logistic Regression using the SAS System. Theory and Application.* Cary, NC, USA: SAS Institute

Altman, E. I., 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, vol. 23, no. 4, pp. 589-609

Arellano, M., 2003. *Panel Data Econometrics. Advanced Texts in Econometrics*. New York, USA: Oxford University Press

Audretsch, D. B., 1991. New-Firm Survival and the Technological Regime. *The Review of Economics and Statistics*, vol. 73, no. 3, pp 441- 450

Audretch, D. B. and T. Mahmood, 1995. New Firm Survival: new Results using a Hazard Function. *The Review of Economic and Statistics*, vol. 77, no. 1, pp. 97-103

Bates, T., 1990. Entrepreneur Human Capital Inputs and Small Business Longevity. *The Review of Economics and Statistics*, vol. 72, no. 4, pp. 551-559

Beaver, W., 1966. Financial Ratios as Predictors of Bankruptcy. *Journal of Accounting Research*, vol. 6, pp. 71-102

Beaver, W., 1968. Market Prices, Financial Ratios, and the Prediction of Failure. *Journal of Accounting Research*, vol. 8, pp. 179-92

Berger, P. G. and W. Ofek, 1995. Diversification's Effect on Firm Value. *Journal of Financial Economics,* vol. 7, pp. 39-65

Borup, L. and M. Lykke, 2003. *New Capital-Adequacy Rules for Credit Institutions.* Monetary Review 3rd Quarter 2003, Danmarks Nationalbank

Brito, P. and A. S. Mello, 1995. Financial Constraints and Firm Post-entry Performance. *International Journal of Industrial Organisation*, vol. 13, pp. 543-565

Bunn, P., 2003. *"Company-accounts-based Modelling of Business Failure.* Financial Stability Review: December 2003, Bank of England

Crowder, M., 2001. *Classical Competing Risks.* London, U.K.: Chapman & Hall/CRC

Danmarks Nationalbank, 2003. *Danish Government Borrowing and Debt 2003.* Copenhagen, Denmark: Danmarks Nationalbank

Danmarks Nationalbank, 2004. *Financial Stability 2004.* Copenhagen, Denmark: Danmarks Nationalbank

Ericson, R. and A. Pakes, 1995. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *Review of Economic Studies*, vol. 62, pp. 53-82

Frame, W. S., A. Srinivasan and L. Woosley, 2001. The Effect of Credit Scoring on Small-business Lending. *Journal of Money, Credit and Banking*, vol. 33, no. 3, pp. 813 - 825

Gort, M. and S. Klepper, 1982. Time Paths in the Diffusion of Product Innovations. *The Economic Journal*, vol. 92, no. 367, pp. 630-653

Greene, W., 2003. *Econometric Analysis.* New Jersey, USA: Prentice Hall

Harhoff, D., K. Stahl and M. Woywode, 1998. Legal Form, Growth and Exit of West German Firms: Empirical Results for Manufacturing, Trade and Service Industries. *Journal of Industrial Economics,* vol. 46, pp. 453-488

Jenkins, S., 1995. Easy Estimation Methods for Discrete-Time Duration Models. *Oxford Bulletin of Economics and Statistics*, vol. 57, no. 1, pp. 129-138

Jenkins, S., 2003. *Survival Analysis.* Unpublished manuscript

Jovanovic, B., 1982. Selection and the Evolution of Industry. *Econometrica*, vol. 50, no. 3, pp. 649-670

Jovanovic, B., 1993. The Diversification of Production. *Brookings Papers on economic activity: Microeconomics*, vol. 1, pp. 197-247

Jovanovic, B. and G. MacDonald, 1994. The Life Cycle of a Competitive Industry. *Journal of Political Economy*, vol. 102, no. 2, pp. 322-347

Kiefer, N. M., 1988. Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, vol. 26, pp. 646-79

Klepper, S., 1996. Entry, Exit, Growth and Innovation over the Product Life Cycle. *American Economic Review,* vol. 86, no. 3, pp. 562-583

Konkurrencestyrelsen, 2003. *Konkurrenceredegørelse 2003.* Copenhagen, Denmark: The Danish Competition Authority

Lamont, O. A. and C. Polk, 2002. Does Diversification Destroy Value? Evidence from the Industry Shocks. *Journal of Financial Economics*, vol. 63, pp. 51-77

Malchow-Møller, N. and M. Svarer, 2002. *Estimation of the Multinomial Logit Model with Random Effects*. Discussion Paper No 2002-16. Accepted for publication in Applied Economic Letters

Møller, M., N. C. Nielsen and J. Ø. Poulsen, 1998. Hvem bærer kreditrisikoen på danske virksomheder? *Nationaløkonomisk Tidsskrift*, vol. 136, pp. 315-332.

Ohlson, J. A. , 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, vol. 19, pp. 109-131

Pakes, A. and R. Ericsson, 1998. Empirical Implications of alternative Models of Firm Dynamics. *Journal of Economic Theory*, vol. 79, pp. 1-45

Rajan, R. , H. Servaes and L. Zingales, 2000. The Cost of Diversity: The Diversification Discount and Inefficient Investment. *The Journal of Finance*, vol. LV, no. 1, pp. 35-80

Statistics Denmark, 2001. *Nye virksomheders overlevelse 1994-1998.* NYT fra Danmarks Statistik nr. 178

Stiglitz, J. E. and A. Weiss, 1981. Credit Rationing in Markets with Imperfect Information. *The American Economic Review*, vol. 71, no. 3, pp. 393-421

Shumway, T., 2001. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, vol. 74, no. 1, pp. 101-124

Tirole, J., 1997. *The Theory of Industrial Organization*. Cambridge (Mass.), USA: The MIT Press

Train, K., 2003. *Discrete Choice Methods with Simulation*. New York, USA: Cambridge University Press

Wooldridge, J. M., 2002. *Econometric Analysis of Cross section and Panel data.* Cambridge (Mass.), USA: The MIT Press

Wooldridge, J. M., 2003. *Introductory Econometrics. A Modern Approach*. Ohio, USA: Thomson South-Western

7. Appendix: Data

Definitions

Return on net assets = primary operating result as a ratio of assets[13]

The solvency ratio = equity capital divided by total liabilities

Short term debt to total assets

Firm size = log(total assets)

Concentration: The concentration in a specific sector is measured by the CR4-index, which is calculated as the sum of the market shares in the four largest companies as a percentage of the total domestic turnover in a specific sector. The index is based on the VAT statistics calculated by Statistics Denmark. For details on the CR4 index, see Konkurrencestyrelsen (2003:chapter 2.3).

Firm age: A dummy for every year is constructed (reference category is firms that are one year old. Firms that are 30 years old or more have the dummy 30 years old or more). Only significant dummies are included and presented in the final estimation results.

Critical comments from the auditors: If there are critical comments from the auditors, the dummy is equal to 1. The critical comments can be "illegal loans have been adopted", "financial statement is incomplete", "inconsistencies in the profit and loss account", etc. These comments indicate illegal activities or that there are discrepancies in the financial statement. Firms with the following critical comments are not included in the "critical comments from the auditors" measure, as they indicate that the firm is not viable (and they are thus likely to be correlated with the dependent variable, the E1 measure): "operation cannot be continued", "there are reservations made to the continuation of operation", etc.

Legal status: If the company is a private limited liability company, the dummy is equal to 1. (Private limited liability companies are compared to public limited liability companies).

Publicly traded company: If the company is publicly traded, the dummy is equal to 1.

---

[13] The primary operating result in year t is divided by total assets in year t. One could also have chosen to divide by the total assets in year t-1 or to take the average of assets in year t and t-1, as it could be argued, that these figures would better reflect the actual resources available to the firm. Nonetheless, assets in year t is used since 1) otherwise, a whole cross-section of observations would be lost, 2) the companies might change the accounting principles as to some extent they can choose between different types of accounting methods, and 3) companies might have acquired new companies or sold of divisions, and therefore that it is not the same company as the year before.

Ownership variables are constructed: Public sector ownership: If the company is owned by the public sector, the dummy is equal to 1. Fund ownership: If the company is owned by a fund, the dummy is equal to 1.

Dummy variables indicating whether or not the specific firm is an ultimate parent company or a wholly-owned subsidiary are constructed: Ultimate parent company: If the company is an ultimate parent company, the dummy is equal to 1. Wholly-owned subsidiary: If the company is a wholly-owned subsidiary, the dummy is equal to 1.

Sector affiliation: There is a dummy for each main category according to table 7.1. Statistics Denmark does not have a sector affiliation category called "IT and telecommunication companies", and so an IT and telecommunication company dummy is constructed. A firm can have 9 sector affiliation codes. If one of the affiliation codes is in the IT and telecommunication sector, which is defined as manufacturing of IT software, hardware, etc., manufacturing of telephone sets, switchboards, and telex apparatus, etc., then the IT and telecommunication dummy is set to 1.

Table 7.1: Sector affiliation

Sector Affiliation	NACE-codes
1. Farming	01
2. Forestry	02
3. Fishing	05
4. Mining	10-14
5. Manufacturing	15-37
6. Energy (“Production of electricity, manufacturing of gas, collection, purification and distribution of water")	40-41
7. Construction (“Construction of buildings and civil engineering works, various contractors and other building completion”)	45
8. Trade and hotel (“Wholesale, retail, repair and hotels”)	50-52, 55
9. Transport	60-64
10. Business service (“Development and selling of real estate, renting, legal activities, advertising, etc."), (except 74.15: non-financial holding companies)	70-74
11. Public service activities (“General (overall) public service activities, education, hospital activities”)	75, 80, 85
12. Organisations, etc. (“Collection and treatment of waste, activities of business and employers organisations, etc., motion picture, video, radio, television, etc., laundering for industrial or commercial clients”)	90-93
13. Not stated	98
14. Unknown*	N.A.

Note: A firm can have up to 9 different sector affiliation codes. The first of these codes, the primary sector affiliation, is used to classify the firms in the various sectors. Unknown* means, that the sector affiliation is not registered in the data base. Reasons for being unknown can be various: The company can be a new company, the company can have exited before 1998 (Up to 1998, The Danish Business Information Bureau deleted the sector affiliation when firms exited the data base)

Diversification: Diversification in 2 sectors (related business): If the company is operating in 2 sectors (within the same main sector), the dummy is equal to 1. Diversification in 3-9 sectors (related business): If the company is operating in 3 – 9 sectors (within the same main sector), the dummy is equal to 1. Diversification in 2 sectors (unrelated business): If the company is operating in 2 sectors (not in the same main sector), the dummy is equal to 1. Diversification in 3-9 sectors (unrelated business): If the company is operating in 3 – 9 sectors (and at least 2 are not in the same main sector), the dummy is equal to 1.

Macroeconomic environment: Controls for the macroeconomic environment are put in (a dummy for every year).

Bank: As a first go on the problem the model was estimated with a dummy equal to 1 if the company has registered a primary bank connection. Otherwise it was zero. The variable was not significant in the estimations. Then the idea was to divide the banks according to the volume of working capital. This would follow the way the Danish Financial Supervisory Authority (FSA) divides the Danish banking sector. The Danish FSA works with the following groups: Banking institutions in category 1 have a working capital of kr. 25 billion and above. Banking institutions in category 2 have a working capital from kr. 3 billion up to kr. 25 billion. Banking institutions in category 3 have a working capital from kr. 250 million up to kr. 3 billion. Banking institutions in category 4 have a working capital of less than kr. 250 million. New estimations were conducted, with the reference dummy being firms that do not register a primary bank connection at all. Compared to firms that do not register a primary bank connection, it was estimated whether or not there was on effect of registering a bank in the Danish FSA's category 1, 2, 3 or 4.

Location: Based on the postal codes, the firms are divided into four groups depending on the location of the firms and the number of inhabitants in the local authorities that the firms belong to. Group 1, 2, 3 and 4 consist of local authorities with 50,000 inhabitants or more, and group 5 consists of local authorities with 50,000 inhabitants or less.[14] Based on these groups, four dummies are constructed (the reference category is equal to group 1). Group 1: Copenhagen and Frederiksberg. Group 2: Local authorities ("kommuner") in the county of Copenhagen ("Københavns Amt") (Ballerup, Brøndby, Dragør, Gentofte, Gladsaxe, Glostrup, Herlev, Albertslund, Hvidovre, Høje-Taastrup, Ledøje-Smørum, Lyngby-Taarbæk, Rødovre, Søllerød, Ishøj, Tårnby, Vallensbæk and Værløse). Group 3: Local authorities in the county of Frederiksborg and Roskilde (Allerød, Birkerød, Farum, Fredensborg-Humlebæk, Frederikssund, Frederiksværk, Græsted-Gilleleje, Helsinge, Helsingør, Hillerød, Hundested, Hørsholm, Jægerspris, Karlebo, Skibby,

[14] In the database the postal code of the firms is included. As the postal codes do not follow the local authorities in all cases, the postal codes have been attached to the local authority which has the largest part of the inhabitants in the postal code area.

Skævinge, Slangerup, Stenløse, Ølstykke, Bramsnæs, Greve, Gundsø, Hvalsø, Køge, Lejre, Ramsø, Roskilde, Skovbo, Solrød and Vallø). Group 4: Other local authorities with 50,000 inhabitants or more ("i største bymæssig bebyggelse") (Odense, Esbjerg, Kolding, Randers, Århus and Ålborg). Group 5: Local authorities with less than 50,000 inhabitants.

Descriptive Statistics

Solvency ratio

Number of firms	Number	Average	St.dev.	Max	Median	Min
E1	2586	0.008075	0.315697	1	0.03895	-0.9938
E2	856	0.510782	0.399105	1	0.5866	-0.9839
E3	1224	0.295536	0.278928	1	0.2659	-0.7865
Active	163684	0.296174	0.236163	1	0.267	-1

Short term debt to total assets

Number of firms	Number	Average	St.dev.	Max	Median	Min
E1	2586	0.82202	0.347881	1.9933	0.81475	0
E2	856	0.435811	0.380759	1.9839	0.33605	0
E3	1224	0.57978	0.290284	1.7865	0.5811	0
Active	163684	0.54654	0.248504	2.1882	0.5371	0

Return on net assets

Number of firms	Number	Average	St.dev.	Max	Median	Min
E1	2586	-0.10053	0.332614	2.4962	-0.0317	-2.9
E2	856	-0.04946	0.428661	2.7472	-0.00735	-2.9806
E3	1224	0.062423	0.217802	1.2439	0.05695	-2.0085
Active	163684	0.084787	0.164132	2.9287	0.0801	-2.9535

Size measured as log(assets)

Number of firms	Number	Average	St.dev.	Max	Median	Min
E1	2586	8.207602	1.278804	14.386	8.11	2.996
E2	856	7.708614	1.473337	13.579	7.683	4.22
E3	1224	9.825529	1.610659	15.801	9.746	5.347
Active	163684	8.924396	1.37991	18.078	8.74	3.178

Age

Number of firms	Number	Average	St.dev.	Max	Median	Min
E1	2586	12.56342	13.15399	209	8	1
E2	856	16.36449	13.6881	142	12	1
E3	1224	22.48366	42.42949	898	15	1
Active	163684	18.70645	20.3886	897	13	1

Sector affiliation

Number of firms	Farming	Forestry	Fishing	Mining	Manufac-turing	Energy	Con-struction
E1	30	1	9	0	595	0	429
E2	4	0	4	1	106	0	107
E3	4	1	9	2	316	9	87
Active	2447	94	696	373	39276	159	26187

Number of firms	Trade etc.	Transport	Busi-ness service	Public service activities	Organi-sations	Not stated	Unknown
E1	626	175	404	23	46	74	174
E2	212	40	175	19	19	46	123
E3	346	78	210	10	34	19	99
Active	52373	9530	23267	4050	2957	1472	803

Critical comments, legal status, listed companies, ultimate parents, wholly owned subsidiaries

Number of firms	Critical comments from the auditors	Legal status: Private limited companies	Listed companies	Ultimate parents	Wholly owned subsidiaries
E1	487	1532	0	49	563
E2	60	368	0	6	328
E3	15	208	0	49	909
Active	7388	65277	547	6388	46899

Location

Number of firms	Group 1	Group 2	Group 3	Group 4	Group 5
E1	415	333	300	423	1115
E2	155	134	102	139	326
E3	170	289	96	219	450
Active	17472	24409	18657	26497	76649

IT-dummy and primary bank

Number of firms	IT dummy	Primary Bank			
		Cat. 1	Cat. 2	Cat. 3	Cat. 4
E1	246	6	101	149	10
E2	52	6	13	21	1
E3	106	5	25	30	0
Active	7388	442	4763	8991	202

Diversification and ownership

Number of firms	Diversification: Same		Diversification: Different		Ownership	
	2 sectors	3-9 sectors	2 sectors	3-9 sectors	Public sector	Fund
E1	327	99	201	99	2	1
E2	86	16	49	15	4	5
E3	200	50	111	86	6	5
Active	31662	12016	15441	10645	262	670

Concentration

Sector affiliation	1995	1996	1997	1998	1999	2000	2001
1. Farming	3.0	3.0	3.2	3.6	3.8	3.3	3.9
2. Forestry	45.3	43.8	42.9	47.6	44.3	41.9	31.9
3. Fishing	4.3	5.2	5.0	5.3	6.0	8.9	7.2
4. Mining	29.1	33.0	52.4	53.8	58.8	59.2	64.9
5. Manufacturing	67.9	67.6	67.9	68.3	68.9	70.9	69.9
6. Energy	44.2	44.5	46.7	48.5	46.8	43.5	43.2
7. Construction	11.5	11.6	11.6	12.4	12.1	14.0	17.2
8. Trade & Hotel	35.3	34.6	35.2	34.2	34.2	35.2	35.8
9. Transport	56.7	57.0	57.3	58.8	60.3	61.7	62.2
10. Business service	49.1	49.2	48.2	48.7	47.7	47.9	48.6
11. Public service	5.7	4.6	3.7	3.8	3.6	3.8	5.4
12. Organisations etc.	41.7	42.2	42.0	43.3	43.9	44.3	43.9
13. Not stated[1]	47.1	38.7	3.2	4.0	5.2	19.1	18.6

Note 1. The figures in category 13 are the ones that The Danish Competition Authority has received from Statistics Denmark. As it is not possible to construct a concentration index in the sector called "unknown", it is chosen to calculate the concentration index for this sector as an (un-weighted) average of the concentration indices in the other sectors.

Source: The Danish Competition Authority and own calculations.

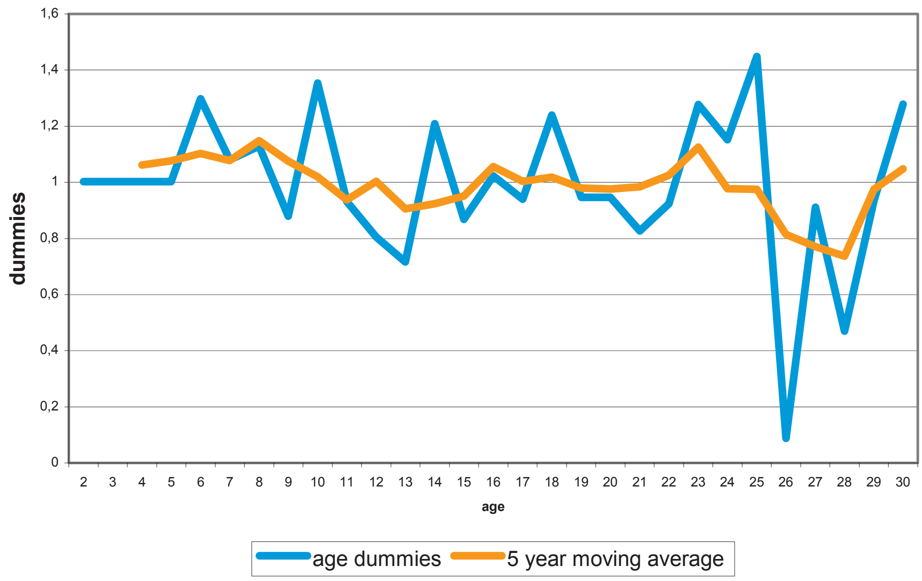
8. Appendix: Figures and Tables

Table 8.a: Competing-risks model: Core variables

Variables	E1	E2	E3
Firm Age (dummies)	See figure 4.2.a.	See figure 8.a below	Only two dummies are significant
Short term debt to total assets	0.4452*	2.1261*	0.7687*
Return on net assets	-1.3417*	-1.1787*	-0.5271*
Solvency ratio	-2.5103*	4.0548*	0.7692*
Firm size	- 0.1180*	-0.6618*	0.2587*

Note: The dummy for publicly traded companies is not included in the estimations as no publicly traded firm enters financial distress. Concerning the controls: there is controlled for the macroeconomic environment and for the various sectors. In the estimations, farming and forestry is included in the same sector affiliation category, as the data were too sparse otherwise. The same is true for mining, energy and construction. The primary bank categories have been altered: As the data was too sparse otherwise, firms that register a bank in category 3 or 4 are now in the same group.
* indicates that the variable is significant at the 1 per cent level.
** indicates that the variable is significant at the 5 per cent level.

Figure 8.a: Competing-risks model: Voluntary exits: Duration dependence



Note: The figure sketches the age dummies (reference dummy is firms that are equal to 1 year old). The last dummy is also called 30 years old or older. All dummies, except one (dummy_age28), are significant at the 5 per cent level. Most dummies are significant at the 1 per cent level.
Note: One big difference between this figure and figure 4.2.a, which pictures the age dummies of the financially distressed firms, is that here all dummies are positive (in figure 4.2.a all dummies are negative).

Table 8.b: Competing-risks model: Proxies

Variables	E1	E2	E3
Owned by the public (dummy)	0.1270 (not sign.)	1.8226*	0.2392 (not sign.)
Owned by a fund (dummy)	-1.4693 (not sign.)	0.2905 (not sign.)	-0.1239 (not sign.)
Diversification 2 sectors (related) (dummy)	-0.3535*	-0.5940*	-0.3654*
Diversification 3–9 sectors (related) (dummy)	-0.3933*	-0.8738*	-0.9830*
Diversification 2 sectors (unrelated) (dummy)	-0.2139*	-0.4347*	-0.3541*
Diversification 3–9 sectors (unrelated) (dummy)	-0.3782*	-0.9257*	-0.5301*
Local authority group 1 (reference dummy)			
Local authority group 2 (dummy)	-0.3240*	-0.1858 (not sign.)	0.2014**
Local authority group 3 (dummy)	-0.2288*	-0.1472 (not sign.)	-0.2628**
Local authority group 4 (dummy)	-0.1195 (not sign)	-0.3088**	0.0862 (not sign.)
Local authority group 5 (dummy)	-0.2010*	-0.3258*	-0.0851 (not sign.)
Concentration	-0.00317 (not sign.)	0.00774 (not sign.)	0.00378 (not sign)
Critical comments from the auditors (dummy)	1.0724*	0.3539**	-0.7454*
Ultimate parent companies (dummy)	0.3767**	-0.2210 (not sign.)	-0.4688*
Wholly-owned subsidiaries (dummy)	-0.3050*	0.4940*	1.5431*
Private limited liability company (dummy)	0.4174*	-0.6769*	-0.3929*
Public limited liability company (reference dummy)			

Note: See the note to table 8.a.

Table 8.c: A comparison: Core variables

Variables	Competing-risks model (E1)	Pooled logit model	E1event model
Firm Age (dummies)	See figure 4.2.a.	Three significant dummies: -0.1244** (age 2 – 18) -0.3524* (age 19) -0.2771* (age 20 and above)	Same shape as figure 4.2.a.
Short term debt to total assets	0.4452*	0.7348*	0.4406*
Return on net Assets	-1.3417*	-1.4931*	-1.2477*
Solvency ratio	-2.5103*	-0.5404*	-2.5878*
Firm size	- 0.1180*	- 0.1129*	-0.1091*

Note: See the note to table 8.a.

Table 8.d: A comparison: Proxies

Variables	Competing-risks model (E1)	Pooled logit model	E1event model
Owned by the public (dummy)	0.1270 (not sign.)	0.9479*	0.0642 (not sign.)
Owned by a fund (dummy)	-1.4693 (not sign.)	-0.2237 (not sign.)	-1.4621 (not sign.)
Diversification 2 sectors (related) (dummy)	-0.3535*	-0.4303*	-0.3472*
Diversification 3–9 sectors (related) (dummy)	-0.3933*	-0.6699*	-0.3894*
Diversification 2 sectors (unrelated) (dummy)	-0.2139*	-0.3152*	-0.2085*
Diversification 3–9 sector (unrelated) (dummy)	-0.3782*	-0.4725*	-0.3753*
Local authority group 1 (reference dummy)			
Local authority group 2 (dummy)	-0.3240*	-0.1386**	-0.3228*
Local authority group 3 (dummy)	-0.2288*	-0.2467*	-0.2224*
Local authority group 4 (dummy)	-0.1195 (not sign)	-0.1287**	-0.1169 (not sign.)
Local authority group 5 (dummy)	-0.2010*	-0.2440*	-0.1961*
Concentration	-0.00317 (not sign.)	-0.00501 (not sign.)	-0.00353 (not sign.)
Critical comments from the auditors (dummy)	1.0724*	-0.8601*	1.0713*
Ultimate parent companies (dummy)	0.3767**	0.0976 (not sign.)	0.3756**
Wholly owned subsidiaries (dummy)	-0.3050*	0.4521*	-0.3370*
Private limited liability company (dummy)	0.4174*	-0.0701 (not sign.)	0.4271*
Public limited liability company (reference dummy)			

Note: See the note to table 8.a.

9. Appendix: Predictions (Competing-risks Model and the E1event Model)

Figure 9.a: Firms predicted as non-events: the distribution of the predictions

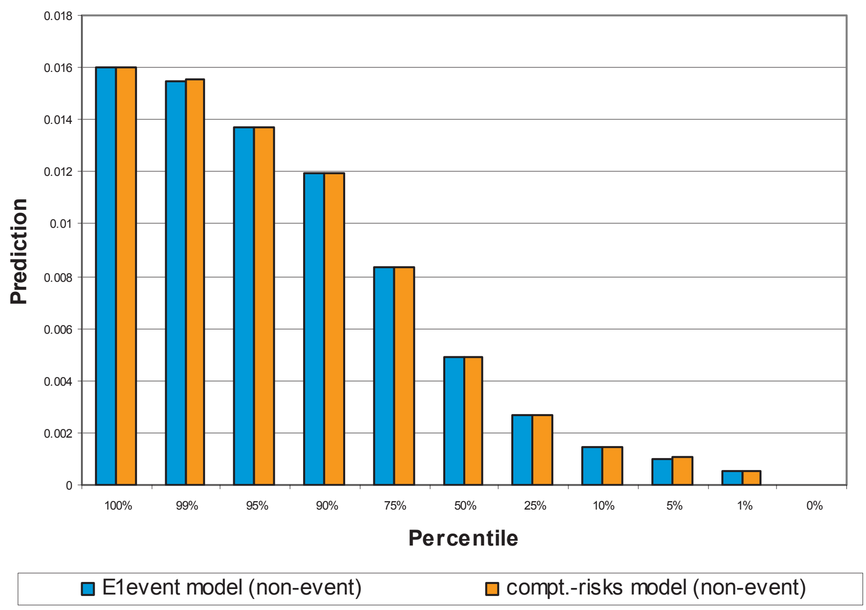


Figure 9.b: Firms predicted as financially distressed (events): the distribution of the predictions