

ANO 2001 / 10

Oslo
December 5, 2001

Working Paper

Financial Analysis and Structure Department

Research Department

A Model of Bankruptcy Prediction

by

Eivind Bernhardsen

*Working papers from Norges Bank can be ordered via the Internet:
www.norges-bank.no/english/publications
or from Norges Bank, Subscription service,
P.O.Box. 1179 Sentrum, 0107 Oslo, Norway.
Tel. +47 22 31 63 83, Fax. +47 22 41 31 05*

Norges Bank's Working papers present research projects and reports (not usually in their final form), and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties.

Views and conclusions expressed in Working Papers are the responsibility of the authors alone.

ISSN 0801-2504
ISBN 82-7553-186-1

A Model of Bankruptcy Prediction

By

Eivind Bernhardsen

December 5, 2001

Abstract

In this thesis, a model of bankruptcy prediction conditional on financial statements is presented. Apart from giving a discussion on the suggested variables the issue of functional form is raised. The specification most commonly applied for the bankruptcy prediction model implies that the rate at which two variables can substitute another holding predicted risk unchanged will be constant. If the aspect captured by single financial ratios is considered less a substitute for any other aspect as this ratio grows, this restriction may not be appropriate. Specifically, the structure of constant compensation will make predictions sensitive to non-credible outliers. A specification of the logit model which allows for flexible rates of compensation is motivated. The model is estimated and the regression results are reported. Second; by questioning the direct connections between financial ratios and the particular outcome of bankruptcy, a model structure which determines an upper bound on probability estimates is explored. By reference to a simple model of misclassification, the specification distinguishes between the probability of bankruptcy and the probability of insolvency. Whereas the predicted probabilities of bankruptcy can be evaluated empirically, the event of insolvency is not observable. Nevertheless; conditional on the model structure, probabilities can be derived for this event as well. An evaluation is given on the ability of the model to measure the over-all development in credit risk for the Norwegian limited liability sector. Individual probabilities of bankruptcy are multiplied with the firms debt to generate a prediction of expected loss in absence of recovered values. This measure is then aggregated and fitted with total loan losses for the Norwegian banking sector over the years 1989-2000. Finally, the possibility of assessing the effect of macro variables in a short panel of firms is explored. With reference to an aggregation property of the probit model, a suggestion is given on how to estimate time-specific effects on aggregate data as a means to identify macro coefficients that can be included in the micro-level model.

Keywords: Bankruptcy, logit analysis, non-linear estimation, aggregation
JEL Classification: G33, C35, C43

Preface:

By this thesis I complete the cand.oecon. degree at the University of Oslo. The thesis was financed by Norges Bank and was written as a part of a project on the development of supervisory credit risk models. Participating on this project was inspiring, both professionally and as a personal experience. I wish to thank the project group; Kai Larsen, Trond Eklund, Terje Lensberg and Elisabeth Axe. I also wish to thank Bjarne Guldbrandsen and Harald Karlsen for useful comments. Kjersti Gro Lindquist and Bent Vale from the Research Department provided me with great support through the process. In particular I wish to thank my supervisor John Dagsvik.

Contents

1	Introduction	1
1.1	The event of bankruptcy	1
1.2	The decision of continuance	1
1.3	Restructuring models	2
1.4	The informational content of the financial statement	2
1.5	Bankruptcy prediction models	4
2	Methodology	6
2.1	The logit and probit models	6
2.2	The random effects probit	7
2.3	The method of maximum likelihood	8
3	A model with flexible rates of compensation	9
3.1	The model	9
3.2	Estimation of the model	14
3.3	The data set	14
3.3.1	A note on sample selection	15
3.3.2	A note on the quality of the data	15
3.3.3	A note on the panel specification	15
3.4	The variables	16
3.5	Model estimates	21
3.6	Predictive ability	24
4	Aggregate predictions	26

5	Financial distress versus bankruptcy	31
5.1	The model	31
5.2	Estimation	33
6	A suggestion on further research	35
6.1	Aggregation in the linear regression model	36
6.2	Aggregation in the probit model	36
6.3	Identification of the micro model coefficients	41
7	Bibliography	42
8	Appendix	44
8.1	Summary statistics	44
8.2	Regression output. Loan losses and risk weighted debt.	46

1 Introduction

1.1 The event of bankruptcy

The Norwegian bankruptcy legislation states that a debtor shall begin bankruptcy proceedings if the debtor is insolvent. The debtor is considered insolvent if he is unable to fulfill his economic obligations as they mature. He is not considered insolvent if his property and income are sufficient to cover the obligations. The Norwegian penal code §283a requires a debtor to petition for bankruptcy when the debtor has reason to believe that the business is run at the expense of the creditors.

Factors which can contribute to the understanding of corporate bankruptcy can be found both in the fields of Economics and in the theory of Business Management. However, the many attempts to specify a model of bankruptcy prediction based on causal specifications of underlying economic determinants has not fully succeeded. The difficulties of merging the theoretical and empirical fields may arise from the diversity of the phenomenon. Firms are heterogenous and the available information is limited. Furthermore, the event of bankruptcy is twofold as the decision of whether or not to continue operations is not directly connected to the particular outcome of bankruptcy. In search of explanatory factors we need not only to identify the factors that influenced on the insufficiency of the firm's performance, but for the firms that do fail we need to explain why the particular outcome of bankruptcy was observed, and not a timely liquidation, a merger, or a restructuring of debt.

1.2 The decision of continuance

If the establishment and abolishment of the firm can be viewed as a reversible investment decision, or the decision cannot be postponed, at any point in time continuance is optimal if the present value of operations is in excess of the liquidation value of the firm. This result is referred to as the standard net present value rule (NPV). If non of the above conditions hold, NPV need not hold and the decision of continuance is better analyzed in

a dynamic framework. The framework of such investment problems is discussed in Dixit and Pindyck (1994). The option to postpone the investment decision will be valuable, and should therefore be priced in the alternative cost. Compared with the NPV, at any point in time a wedge is added to the critical levels of the decision rules. This result is indeed relevant for the decision of firm continuance: If the entry or exit of markets are sufficiently costly and the variance of outcomes sufficiently high the firm may choose to operate even at a negative contribution margin.

1.3 Restructuring models

In presence of a positive probability of bankruptcy, the value of a company can be viewed as a call option which will be valuable to the shareholders only if the market value of the company is considered greater than the company debt at the date of maturity. If the option is "out of money" the creditors will have to bear the loss (i.e. a bankruptcy petition is filed). The call option need not be exercised, and thus there will be an asymmetry in the risk faced by shareholders and creditors. This asymmetry may cause the troubled firm to engage in particular risky projects in effort to recover some value, and so there is a potential for inefficient investment decisions. Models of debt restructuring¹ emphasize the fact that shareholders, bondholders and debtholders will have different priorities on assets liquidated, different ability to control the firm, and different exposure to the risk associated with continuance. By considering different assumptions concerning the underlying setting the restructuring models seeks to analyze what is likely to determine the destiny of a troubled firm.

1.4 The informational content of the financial statement

The financial statement is a filtered representation of information. Decisions are made concerning the classification of income and expenses, the timing of income and expenses

¹Restructuring models are analysed in several studies. See for example Myers(1977), Bulow and Shoven(1978), and Chen, Weston and Altman (1995).

as well as the valuation of assets and conventions of depreciation. In many cases the firm will have incentives to bias the entries; income tax, profit related pay and debt covenant restrictions are explicitly dependent on the reported figures. The firm may signal profits to attract investors or to win time in a situation of financial distress.

The use of financial ratios to make qualitative statements about the going concern of the firm has a long tradition. However, the generality of constructed ratios are controversial. Any textbook of accounting will emphasize the fact that benchmark values are not directly comparable over different industries. Financial ratios must thus be evaluated in conjunction with additional information related to the nature of the firm and the market in which it operates: Differences in trading cycles and degree of capital turnover, market competition, volatility of revenues and costs and the industry's dependency on the business-cycle are factors of importance.

Moreover; measuring financial ratios is not equivalent with observing "real characteristics", but should rather be considered as "surrogate measures" of the relevant aspects. As emphasized by Morris (1989): A unique economic event can result in a variety of ratio patterns, and a single pattern of ratios can be the result of a variety of underlying economic conditions (fig 1). The business analyst put on the task of giving a subjective evaluation of a firm will therefore use the collection ratios interactively. Different constellations of the financial entries can give rise to hypothesis of the underlying economic conditions. Ideally, the analysis is combined with external sources of information so that an over all profile of the firm can be drawn.

Any statistically derived bankruptcy prediction model implicitly assumes that benchmarking financial ratios makes sense. The limited success of bankruptcy prediction models must be viewed in this perspective. Nevertheless, in practice, bankruptcy prediction models are found useful: The holder of a large portfolio of claims may find it costly to supervise individual developments and therefore use the credit risk model as a means to make a first selection of "follow-ups". Furthermore the "objectiveness" of the statistical model may be appreciated. The ability to discriminate by subjective judgements will potentially depend grossly on who is making the analysis and his current orientation towards general

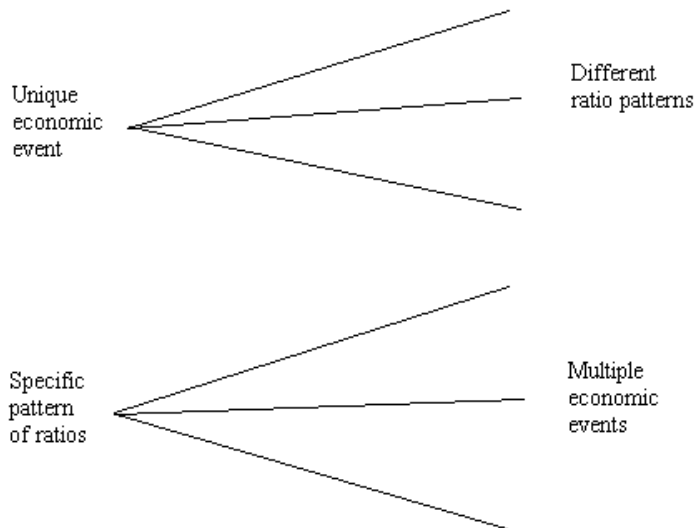


Figure 1: The ambiguity of ratio patterns

economic developments. Even if subjective predictions on average are more effective than those of the statistical model, this source of uncertainty may not be appreciated. If the degree of accuracy of the statistical model can be accurately measured, the model will be particularly useful.

1.5 Bankruptcy prediction models

The study of Beaver (1966) is considered the pioneering work on bankruptcy prediction models. Beaver motivated his model by a framework quite similar to the model of the gamblers ruin². The firm is viewed as a "reservoir of liquid assets, which is supplied by

²In the gamblers ruin model one assumes that net assets follows a random walk process with some fixed probability of a negative cashflow each period. In the case of no access to external capital, the model is quite simple: For a sufficiently long sequence of periods there is always some probability for a clustering of negative cash flows so that the net assets eventually takes on a negative value. For an application of the gamblers ruin, see Wilcox(1976)

inflows and drained by outflows. (...) The solvency of the firm can be defined in 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". By this framework beaver state four propositions:

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

Beaver identified 30 ratios that were expected to capture relevant aspects. By a univariate discriminant analysis, these ratios were applied on 79 pairs of bankrupt/non-bankrupt firms. The best discriminators were "working capital funds flow/total assets" and "net income/total assets" which correctly identified 90% and 88% of the cases.

Altman(1968) conducted a similar study applying multivariate discriminant analysis using the 7 ratios; return on assets, stability of earnings, debt service, cumulative profitability, liquidity, capitalization and size. Applied on 33 pairs of bankrupt/non-bankrupt firms the model correctly identifies 90% of the cases one year prior to failure.

Ohlson(1980) is the first to apply the logit analysis on the problem of bankruptcy prediction. By using 105 bankrupt and 2,058 non-bankrupt firms he is also the first to apply a representative sample. He states that predictive power appears to be less than reported in previous studies.

Recent years, much attention is given to the choice of methodology. Methods like recursive partitioning, neural networks and genetic programming are commonly applied on the bankruptcy prediction problem. Morris (1998) gives a survey on both new and traditional approaches to bankruptcy prediction.

2 Methodology

2.1 The logit and probit models

Assume that the variable $y_i \in \{0,1\}$ is related to an unobservable index y_i^* by a linear function of the explanatory variables $x_{i1}, x_{i2}, \dots, x_{ik}$ and the random term u_i such that:

$$\begin{aligned} y_i^* &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i \\ y_i &= 1 \text{ if } y_i^* > 0 \\ y_i &= 0 \text{ else} \end{aligned} \tag{1}$$

By this structure we have;

$$\begin{aligned} P(y_i = 1 | \beta' x_i) &= P(u_i > -\beta' x_i) \\ &= 1 - F(-\beta' x_i) \end{aligned} \tag{2}$$

where $F()$ is the cumulative distribution function for u . Most commonly u is assumed normally or logistically distributed. If u is assumed normally distributed;

$$F(-\beta' x_i) = \int_{-\infty}^{-\beta' x_i / \sigma} \frac{1}{(2\pi)^{0.5}} \exp\left(-\frac{t^2}{2}\right) dt \tag{3}$$

and the model given by eq(1) is referred to as the probit model. If u is assumed logistically distributed

$$F(-\beta' x_i) = \frac{1}{1 + \exp(\beta' x_i)} \tag{4}$$

and the model of eq(1) is referred to as the logit. The logistic distribution differ from

the normal distribution primarily by being slightly thicker at the tails. The predicted probabilities will be quite similar unless the sample is large and enriched with observations at the tails³.

2.2 The random effects probit

If a panel is constructed one may not want to impose the restriction that u is identically distributed over observations. In the random effects probit model the error term for individual i at time t is decomposed such that;

$$u_{it} = \mu_i + \varepsilon_{it} \quad (5)$$

$$\mu_i \sim N(0, \sigma_\mu)$$

$$\varepsilon_{it} \sim N(0, \sigma_\varepsilon)$$

$$\text{cov}(\mu_i, \varepsilon_{it}) = \text{cov}(\mu_i, x) = \text{cov}(\varepsilon_{it}, x) = \text{cov}(\varepsilon_{it}, \varepsilon_{ir}) = 0$$

$$i = 1, 2, \dots, N \quad (6)$$

$$t, r = 1, 2, \dots, T, t \neq r \quad (7)$$

Thus for each individual the element μ_i is drawn once and added to the constant term. Defining $\sigma = \frac{\sigma_\mu}{\sigma_\varepsilon}$ such that $\rho = \frac{\sigma}{\sigma+1}$ gives the proportion of the total variance contributed by the panel level variance component then;

$$P(y_i|x_i) = \int_{-\infty}^{\infty} \frac{e^{-\frac{\mu_i^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \left\{ \prod_{t=1}^{n_i} F(\beta'x_{it} + \mu_i) \right\} d\mu_i \quad (8)$$

$$F(\beta'x_{it} + \mu_i) = \Phi(\beta'x_{it} + \mu_i) \quad \text{if } y_{it} = 1 \quad (9)$$

$$F(\beta'x_{it} + \mu_i) = 1 - \Phi(\beta'x_{it} + \mu_i) \quad \text{otherwise} \quad (10)$$

The hypothesis of $\rho = 0$ (no individual heterogeneity) can be tested by the likelihood ratio test. By the restriction $\rho = 0$ we get the pooled probit model. If $\rho > 0$ and the

³An introduction to the logit and probit models as well as the method of maximum likelihood are given in Maddala (1983)

pooled probit is estimated, the estimator of β will still be consistent although inefficient. Estimated standard errors of β will be biased. However, this bias can (inefficiently) be adjusted for by summing within individuals when calculating the standard errors⁴.

2.3 The method of maximum likelihood

The model of eq(1) can be estimated by the method of maximum likelihood. By assuming that u_i , $i = 1, 2, \dots, N$ is independently distributed the joint probability of observing $\{y_1, y_2, \dots, y_N\}$, each element respectively conditional on $\{\beta'x_1, \beta'x_2, \dots, \beta'x_N\}$, is given by the likelihood function:

$$L = \prod_{i=1}^N F(\beta'x_i)^{1-y_i} (1 - F(\beta'x_i))^{y_i} \quad (11)$$

Maximizing eq(11) with respect to the coefficient vector β is equivalent to maximizing the log of L and is solved by setting

$$\frac{\partial \ln L}{\partial \beta_j} = 0, \quad j = 1, 2, \dots, k \quad (12)$$

Eq(12) will constitute k non-linear equations and must be solved numerically by an iterative procedure. For the logit and probit models the information matrix given by $\frac{\partial^2 \ln L}{\partial \beta \partial \beta'}$ will be positive definite at any stage of the iteration procedure, and the iterations will converge to a maximum of the likelihood function independently of the initial values of β .

⁴A discussion on the random effects probit model is given in Maddala (1987), and Guilkey and Murphy (1992).

3 A model with flexible rates of compensation

3.1 The model

Applying the specification of (1) for the bankruptcy prediction model, the explanatory variables $x_{i1}, x_{i2}, \dots, x_{ik}$ would be financial ratios computed from the balance sheet, and $\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} = \beta' x_i$ is a measure of the financial soundness of the firm. If this measure exceeds a critical value, the firm is assumed to go bankrupt. The critical value of this variable is assumed to vary among individuals and thus the stochastic term u_i is introduced⁵. By any choice of a monotonic distribution function for u this structure will imply a constant rate of compensation between variables: The change (in units) in x_r needed to offset a marginal increase in the variable x_s such that risk is held unchanged is independent of the values of x_r and x_s . By the logit model we have;

$$P(y = 1 | \beta' x) = \frac{1}{1 + \exp(-\beta' x)} \quad (13)$$

The marginal effect of variable x_r is given by;

$$\frac{\partial P}{\partial x_r} = P(1 - P)\beta_r \quad (14)$$

The marginal effect of x_r is thus implicitly dependent on x_r by the term $P(1 - P)$. The rate of compensation between x_r and x_s is given by

$$\frac{\partial x_r}{\partial x_s |_{dp=0}} = -\frac{\frac{\partial P}{\partial x_s}}{\frac{\partial P}{\partial x_r}} = -\frac{\beta_s}{\beta_r} \quad (15)$$

⁵In the following I will suppress the subscript for individuals. The notation x_k will refer to the variable k for any individual i .

In Laitinen and Laitinen (2000) the appropriateness of applying this specification of the bankruptcy prediction model is questioned. The issue is motivated by the following numerical example:

Consider first a firm for which the ratios "cash to total assets" (measuring liquidity) and "cash-flow to total assets" (measuring profitability) both are at a 5% level. Assume furthermore that at these levels the liquidity measure is considered slightly more critical in such a way that the firm would be considered equally risky at a liquidity level of 3% if profitability doubled to 10%. Thus;

$$\frac{\partial (\text{cash-flow to total assets})}{\partial (\text{cash to total assets})} \Big|_{dp=0} = -\frac{2}{5} \quad (16)$$

Consider next a firm with profitability at 5%, and liquidity measured at 50%. If the same rate of compensation is imposed, a fall in liquidity to 48% would still require a doubling of the profitability measure if predicted risk is to be held unchanged.

Laitinen et. al. argues that one would not be greatly concerned whether liquidity is measured at 50% or 48% and thus a constant rate a of compensation appears unreasonable. The objection could be interpreted in two ways:

- In terms of insolvency risk, the aspect captured by the variable x_r is less a substitute for the aspect captured by variable x_s as the variable x_r increases.
- The variable x_r is considered less likely to be a relevant measure of the target aspect as the variable is measured at more "extreme" values.

By the first suggestion we would like the marginal effect of the variable x_r to decline with the level of x_r at any given P . If the second suggestion is considered relevant we would generally like the marginal effect to decline as the variable deviates from some critical value. Indeed this a relevant issue. Financial ratios are artificially generated by the division of entries like "total assets" or "revenue from operations" as a means to adjust for size or the level of activity. These are rough measures, and will be highly irrelevant for

some firms. If the aspect is irrelevantly measured the variable is likely to take odd values. Most studies conducted on large samples typically apply truncations of the variables at some quantile. However, the significance of the variable will potentially depend on the choice of truncation and thus make it difficult to determine which set of ratios to include in the model. Furthermore, the optimal truncation of one variable may depend on what truncations are applied on the other variables.

Laitinen et. al. does not impose a specific structure, but rather leaves the question of functional form open. A Taylor expansion of the underlying functional relationship at the mean values of the variables is used to motivate the inclusion of cross products and squares of variables in the logit model.

An alternative to this approach would be to impose a specific functional form with the desired properties. If interaction effects between variables are ignored we could consider the model;

$$y^* = \beta_0 + \beta_1 T_1(x_1) + \beta_2 T_2(x_2) + \dots + \beta_k T_k(x_k) + u \quad (17)$$

$$y = 1 \text{ if } y^* > 0 \quad (18)$$

$$y = 0 \text{ else} \quad (19)$$

,where the function $T_r(x_r)$ is possibly non-linear in the explanatory variable x_r . One suggestion for $T_r(x_r)$ would be to apply the Box-Cox transformation;

$$T_r(x_r) = \frac{(x_r)^{\lambda_r} - 1}{\lambda_r} \quad (20)$$

For λ equal to one the transformation is linear. For $\lambda_r < 1$, the transformation will be concave and thus at any level of P the marginal effect of variable x_r will decrease with the level of x_r . If $\lambda_r > 1$ the transformation is convex. By applying this transformation, if both λ_r and λ_s is found not to be significantly different from one, this would suggest that the rate of compensation between x_r and x_s is constant. However, the Box-Cox transformation is not defined for negative values of x . This will be problematic for the

bankruptcy prediction model as the cash-flow measure potentially will take on negative values. Furthermore, assuming a concave relationship, at any given level of P the marginal effect of x_r will be greatest at $x_r = 0$. More generally we might like the marginal effect to decline as x_r deviates from some constant α_r . If we impose the cumulative logistic function for $T_r(x_r)$, by estimating the scale and location parameters specific for each variable we will have this structure:

$$T_r(x_r) = \frac{1}{1 + e^{-\frac{x_r - \alpha_r}{\delta_r}}} \quad , i = 1, 2, \dots, k \quad (21)$$

$\alpha_r = \textit{location parameter for variable } x_r$

$\delta_r = \textit{scale parameter for variable } x_r$

This transformation does not include the linear function as a linear case but on a given interval for x_r , the function $T_r(x_r)$ will be approximately linear if the scale parameter δ_r is sufficiently large. To illustrate this, a plot of $T_r(x_r)$ for $x_r \in [-100, 100]$, $\alpha_r = 0$, and $\delta_r = 10$, $\delta_r = 10$ and $\delta_r = 100$ respectively, is given in fig2.

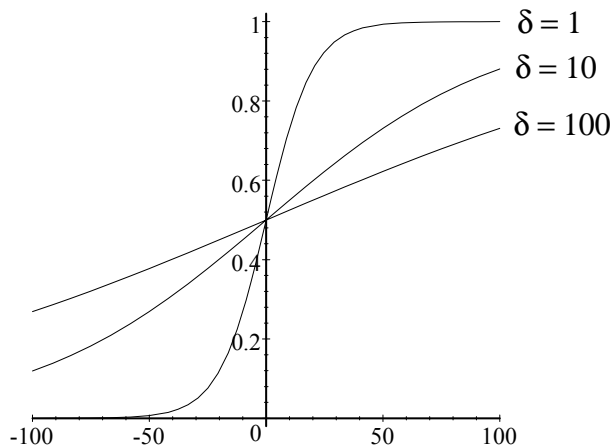


Fig 2: $T(x)$ with various scaling

For a sufficiently large δ_r , the transformation can roughly be viewed as a re-scaling of

the β -coefficient⁶;

$$\begin{aligned} T_r(x_r) &\approx a_r + b_r x_r \\ \Rightarrow \frac{\partial P}{\partial x_r} &\approx P(1 - P)b_r \beta_r \end{aligned} \quad (22)$$

For sufficiently large scale parameters δ_r and δ_s the rate of compensation between the variable x_r and variable x_s will thus be constant.

In general, the rate of compensation between the variable x_r and variable x_s will not be constant, as the marginal effect of the variable x_r will depend explicitly on the value of x_r :

$$\frac{\partial P}{\partial x_r} = P(1 - P)T_r(x_r)(1 - T_r(x_r))\frac{\beta_r}{\delta_r} \quad (23)$$

When $x_r = \alpha_r$, $T_r(x_r)$ will equal 0.5 and thus $T_r(x_r)(1 - T_r(x_r))$ will be at its maximum. Independently of δ_r , the position parameter α_r thus determines which value of x_r maximize $\frac{\partial P}{\partial x_r}$ for a given probability P . As the variable x_r deviates from α_r , $T_r(x_r)(1 - T_r(x_r))$ will approach zero. The rate of compensation between x_r and x_s is given by:

$$\frac{\partial x_r}{\partial x_s |_{dp=0}} = -\frac{\beta_s T_s(x_s)(1 - T_s(x_s)) \delta_r}{\beta_r T_r(x_r)(1 - T_r(x_r)) \delta_s} \equiv g_{rs}(x_r, x_s) \frac{\beta_s \delta_r}{\beta_r \delta_s} \quad (24)$$

For given a given value of $\frac{\beta_s \delta_r}{\beta_r \delta_s}$, the change in x_r needed to compensate a rise in x_s will be larger the more x_r deviates from α_r , and smaller the more x_s deviates from α_s .

⁶In this case, only the product $b\beta$ and not the β coefficient will be identifiable, . Furthermore, it will not be possible to separate a from the constant term of the model.

$$a_r = \frac{1}{1 + e^{\frac{\alpha_r}{\delta_r}}}, \quad b_i = \frac{1}{\left(1 + e^{\frac{\alpha_r}{\delta_r}}\right)^2} \frac{e^{\frac{\alpha_r}{\delta_r}}}{\delta_r}$$

3.2 Estimation of the model

The model can be estimated by the method of maximum likelihood. Conditional on an objective set of initial values of (α, δ) one can switch between estimating β conditional on (α, δ) and (α, δ) conditional on β . By setting sufficiently large initial values of δ the procedure will have the approximately linear transformations as a starting point.

3.3 The data set

The data used in this study was constructed by the SEBRA-database at Norges Bank. The database contains the annual financial statements of all limited liability firms registered at the Norwegian register for business enterprises over the years 1988-1999. The bankruptcy data is computed by Dun and Bradstreet, and is more or less complete from 1990-1999. In the preliminary examination of the data it was found that for most bankrupt firms, there existed a substantial lag between the date of the last registered financial statement and the date of bankruptcy: If the last registered statement were recorded in year t , only 25% of the bankrupt firms are declared bankrupt in year $t+1$, 55% in year $t+2$, and 20% in year $t+3$. Because of this feature of the data, it was decided to use only the years 1990-1996 for estimation. Furthermore it was considered most appropriate to define the endogenous variable by the event "the firm was registered bankrupt within 3 years *and* this year constitutes the last registered financial statement". By this approach a pooled panel structure could be estimated without multiple counting of the responsive event.

Examining this sub-sample some observations were excluded due to a missing bankruptcy variable. Furthermore, firms for which the book value of total assets did not exceed 250.000 NOK were excluded: For these firms the entries of the financial statements were frequently considered difficult to interpret and thus suspected to be plagued with errors of registration. The estimation sample was constructed by the remaining sample, now containing 398.689 observations including 8.436 bankruptcies.

3.3.1 A note on sample selection

The number of registered bankruptcies in the SEBRA database is far less than the number found in the official statistics on the Norwegian limited liability sector⁷. The SEBRA database only includes firms for which the financial statement some year was approved by the Register for Business Enterprises⁸. If the financial statement of a newly established firm is more likely to be disapproved when the bankruptcy risk is high, this will generate a sample selection problem.

3.3.2 A note on the quality of the data

Quite frequently, firms were found to be temporarily absent from the data-set, and the number of firms absent showed significant variation over the estimation period. Furthermore; the year 1994 surprisingly contained a very small number of new establishments (about a tenth of the sample average).

The change in the proportion of bankruptcies recorded in the SEBRA database did not show strict correspondence with official statistics, sometimes not even in signs. The bankruptcy data was gathered from a different source than the financial statements, and the quality of this variable is suspected vary over the estimation period.

The financial statements recorded in the SEBRA-database are adjusted prior to the year of 1992, as an attempt to incorporate the effects of the 1992 Norwegian tax-reform.

The risk of adverse effects due to time-specific sample features was believed to be substantial. A pooling of the data as a means to smooth the sample was therefore preferred.

3.3.3 A note on the panel specification

Tentatively, the random effects probit was estimated. By the likelihood ratio test the restriction $\rho = 0$ could however not be rejected. This was not taken as evidence of

⁷The number of bankruptcies recorded in the SEBRA database was compared to the official numbers of Statistics Norway. On average the number recorded in the database is lower by 30%.

⁸Foretaksregisteret Brønnøysund

absence of individual heterogeneity, but rather as a result of the consistency property of the pooled probit specification.

Considerable effort was made to explore whether a dynamic specification could show useful. However; even for the sample where only firms that were present at $t - 1$ were included, lagged variables and lagged probability predictions (quite surprisingly) showed little significance. Some success was found for dummy variables that captured events like "revenues did drop more than 20% and short term debt did rise". However, the success was limited and it was not considered practical to include these variables.

3.4 The variables

- Liquidity:

$$lik = \frac{\text{Cash and deposits} - \text{Value of short term debt}}{\text{Revenue from operations}}$$

$$ube = \frac{\text{Outstanding payments of public dues}}{\text{Total assets}}$$

$$lev = \frac{\text{Trade creditors}}{\text{Total assets}}$$

- Profitability:

$$tkr = \frac{\text{Result before extra ordinary items} + \text{Ordinary write offs} + \text{Depreciation} - \text{Taxes}}{\text{Total assets}}$$

- Solidity:

$$eka = \frac{\text{Book value of equity}}{\text{Total assets}}$$

$taptek$ = "Current book value of equity is less than the value of equity injected" (dummy)

div = "Dividends paid current year" (dummy)

- Age:

a_x = "Number of years since incorporation ", $x = 1, \dots, 8$ (dummies)

- Size:

$$size = (\ln(\text{Total assets}) - 8.000)^2$$

- Industry characteristics⁹:

$meanek$ = Mean value of the variable eka

$meanlev$ = Mean value of the variable lev

$sdtkr$ = Variance of the variable tkr

The list of explanatory variables applied must be viewed as a suggestion. The variables named " lik ", " tkr " and " eka " are traditionally used for the analysis of credit risk at Norges Bank. In this thesis these variables are used as core measures of liquidity, profitability and solidity. The remaining variables were found by trial and error¹⁰. By the number of observations in the estimation sample, one would expect that some generality can be assumed for these variables. The comment on these variables should however be viewed as suggestive.

- Liquidity:

The amount of cash the firm needs to service its going expenditures will depend fundamentally on the nature of activities, and one should be reluctant to consider benchmark values for liquidity ratios. However; commonly firms are drained in terms of liquid assets immediately prior to bankruptcy, and it may borrow heavily to manage its short term obligations. Commonly, ratios like "short term debt to revenue from operations" and "cash to total assets" are found useful in bankruptcy prediction models. In the credit risk model of Norges Bank the aspect of liquidity is sought captured by the variable "cash minus short term debt to revenue from operations". Applying this variable is analogous with the inclusion of both "short term debt to revenue from operations" and "cash to revenue from operations" if a coefficient restriction is imposed¹¹.

⁹The 5-digit industry code of Statistics Norway was used. The degree of crudnes of this clasification was determined as to include at least 1000 firms.

¹⁰Summary statistics on all variables are reproduced in the appendix.

¹¹Empirical support was not found for this restriction. The restiction was however not found to significantly affect the predictive power of the model, and was applied mainly for practical reasons.

- Profitability:

The profitability of the firm should be considered the driving factor for both the liquidity and solidity aspect. In the long run, the firm must generate a sufficient margin on its operations to be able to service its debt. Sustained negative profits will quickly drain the solidity of the firm, and if the firm is to expand it may need to retain earnings in excess of existing requirements. In the short run, negative profits will quickly drain the liquidity of the firm. Furthermore; the profitability of the firm is likely to influence the ability of obtaining external finance. The aspect of profitability is sought captured by a straight forward measure of return on capital employed.

- Solidity:

If markets are not perfect, the capital structure will be of importance for the contractual relationship between shareholders and debtholders. The greater the share of shareholders equity, the lower the financial risk, and the firm is more likely to obtain external finance. The book value of equity is a residual measure in the balance sheet, and thus directly related to the valuation of the firms assets. Furthermore; the equity share of total assets will give information on the historic performance, and serve as a buffer on future negative profits.

- Outstanding public dues to total assets:

Often bankruptcy proceedings will be initiated by a bankruptcy petition submitted by the revenue authorities. The authorities have definitive procedures for treating default payments on taxes and dues, and will generally not negotiate with an insolvent debtor. It is reasonable to expect that the debtor will give priority to these obligations. Thus, if public dues are used as a liquidity buffer the firm is likely to be in severe distress.

- Trade creditors to total assets:

The "natural" level of trade creditors to total assets will vary extensively among industries, and thus any effort to benchmark this variable is controversial. However, by including both an industry variable (see below) and an individual variable that seek to capture this aspect one can hope to establish whether trade creditors is used as a buffer on liquidity.

- " Book value of equity is less than injected equity"

This variable may indicate to what extent a given level of equity to total assets is the result of accumulated earnings.

- "Dividends are paid current year"

Dividends may be used to signal profitability or, if the firm is troubled, as a means to withdraw assets from the creditors. For these reasons one can easily question the usefulness of including this variable in a model of bankruptcy prediction. However, the Norwegian legislation on limited liability companies states that dividends are not to be paid if there is reason to believe that the firm is in risk of immediate insolvency. If the legislation is obeyed, the variable should serve as a signal of solidity.

- Industry mean of equity to total assets:

If the variable "equity to total assets" is most properly measured by its deviance from the industry mean this variable will show significant. If *only* deviance from industry mean matters, the coefficient on this variable should have the opposite sign of the coefficient on equity to total capital. However, the variable was found negatively correlated with bankruptcy. Accordingly this variable contributes with some additional information concerning the risk related to the industry. One should be careful to give definitive interpretations of this result. However, since solidity is partly a result of retained earnings, one could suspect that industries characterized with high leverage are subject to more competition than industries with low leverage.

In an industry with a high degree of competition we would expect that both entry and exit rates are likely to be higher.

- Industry mean of trade creditors to total assets.

If trade creditors to total assets is most properly measured by its deviance from industry mean, the variable will show statistically significant. However; the variable was found positively correlated with bankruptcy and does thus appear to give additional information concerning the risk of the industry. The result may capture the fact that restaurants and retail business are associated with both a high level of trade creditors and high bankruptcy rates. A dummy variable for restaurants could not compensate for the exclusion of this variable.

- Industry variance of return on capital employed.

Economic risk will be reflected in the variability of a company's earnings over time. If the industry is associated with a high level of variability of earnings and thus a high level of risk, we would expect both a higher rate of prosperous firms as well as a higher rate of bankruptcies.

- Number of years since establishment (dummy variables for each of the first 8 years).

Uncertainty concerning the true costs of production as well as factors concerning the competitive setting makes of establishment of a business risky. Furthermore; the firm may need time to develop a functional organizational structure and sufficient management skills.

- Size of the firm:

The size of the firm is commonly identified as a significant factor in bankruptcy prediction models. Commonly the logarithm of total assets is employed. In this study this variable was not found significant. However the square deviance from 2 mill NOK did enter significantly. This result may indicate that if the firm is sufficiently small, (administrative) bankruptcy costs will exceed the expected liquidation value of the firm, and thus the creditor may not want to initiate bankruptcy proceedings.

3.5 Model estimates

When estimating the model initial values were respectively set such that $\alpha_r = 0$, $\delta_r = 1$, $r = 1, 2, \dots, k$ (Model A) and $\alpha_r = 0$, $\delta_r = 100$, $r = 1, 2, \dots, k$ (Model B). As the variables were measured in percent, the interval $[-100, 100]$ captured at least 98 percent of the observations for any variable. With reference to fig 2; by setting $\delta_r = 100$ the variables will enter approximately linearly in eq(17) at the start of the iteration process.

Table 1. Model(A) estimates:

variable	β	<i>s.e</i>	$\frac{\alpha}{\delta}$	<i>s.e.</i>	$\frac{1}{\delta}$	<i>s.e</i>
eka	-1.4459	0.0604	0.4464	0.0977	0.0782	0.0049
tkr	-1.0948	0.0386	0.1216	0.1274	0.2096	0.0190
lik	-1.4925	0.0421	-2.9618	0.1977	0.1529	0.0087
lev	0.4968	0.0486	1.5224	0.4142	0.2895	0.0660
ube	6.8069	0.2019	-1.1474	0.0243	0.0362	0.0017
a1	0.8380	0.0438
a2	0.9707	0.0382
a3	0.8310	0.0398
a4	0.6729	0.0429
a5	0.5282	0.0468
a6	0.3189	0.0528
a7	0.2689	0.0575
a8	0.2076	0.0638
div	-1.0639	0.0742
taptek	0.5386	0.0419
size	-0.0543	0.0064
meanlev	1.0404	0.1692
meanek	-3.9690	0.2273
sdtkr	1.8229	0.3319
const	-7.0131	0.2786

LR chi(19) = 21909.7

log-L= -29917.726

Table 2. Model(B) estimates:

variable	β	<i>s.e.</i>	$\frac{\alpha}{\delta}$	<i>s.e.</i>	$\frac{1}{\delta}$	<i>s.e.</i>
eka	-1.4006	0.0589	0.3594	0.0987	0.0791	0.0051
tkr	-1.0594	0.0381	-0.0188	0.1301	0.2115	0.0195
lik	-1.3629	0.0391	-3.5081	0.2447	0.1704	0.0107
lev	1.4287	0.1182	-0.5975	0.1169	0.1670	0.0286
ube	2.7226	0.0812	0.7715	0.0418	0.0389	0.0016
a1	0.8297	0.0438
a2	0.9735	0.0382
a3	0.8310	0.0398
a4	0.6753	0.0430
a5	0.5312	0.0469
a6	0.3204	0.0528
a7	0.2709	0.0575
a8	0.2083	0.0639
div	-1.0826	0.0742
taptek	0.5496	0.0419
size	-0.0573	0.0064
meanlev	1.2807	0.1690
meanek	-3.9623	0.2280
sdtkr	1.8229	0.3323
const	-3.6069	0.2458

LR chi(19) = 21846.80

log-L = -29932.503

Analytically I have not explored whether the maximum likelihood problem of determining scale and position parameters conditional on the β estimates has a global maximum independent of the initial values. The estimates differ slightly, but the model estimates can not be considered crucially dependent on the choice of initial values.

The variables "equity to total assets" and "public dues to total assets" enter quite linearly in eq(17) with the scale parameters 12.642 and 41.667 respectively. In contrast, the profitability variable enters with $\delta = 4.7281$. In Fig 3, marginal effects are simulated for Model (A). The simulation is conducted by setting all variables (except for the one

plotted) at their mean values. The simulated firm has been operating for 3 years, did not experience loss in equity and did not pay dividends the relevant year. The profitability variable is effective in a quite narrow interval only. In fact; the model was found to perform surprisingly well when this variable was replaced with a dummy variable for negative profits.

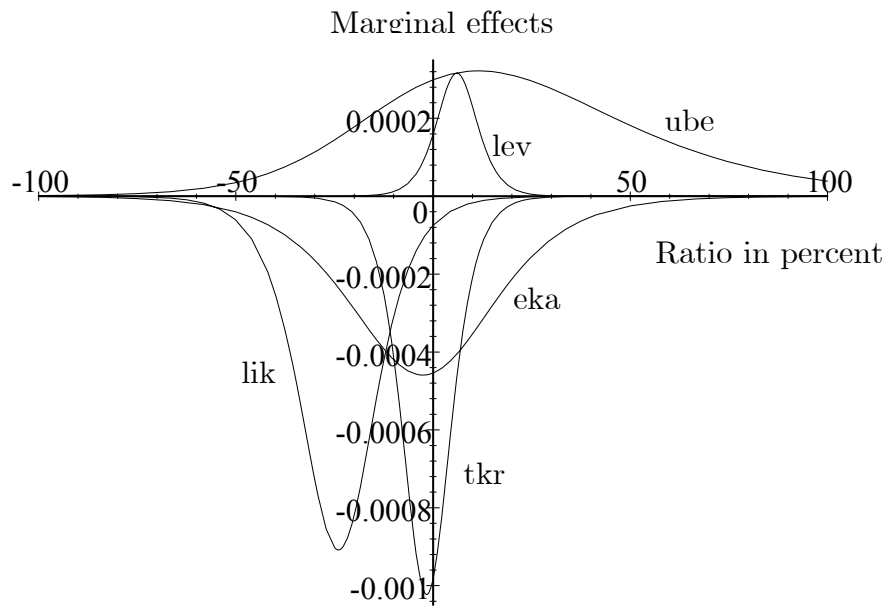


Figure 3. Simulated marginal effects on the probability of bankruptcy.

The rates of compensation between various variables are explored in fig 4. The figure to the left gives the rate at which a fall in solidity compensates a marginal rise in liquidity. The figure to the right gives the rate at which a fall in liquidity compensates a marginal rise in profitability. In the saddle point the marginal effects of both variables are maximized at any given probability p .

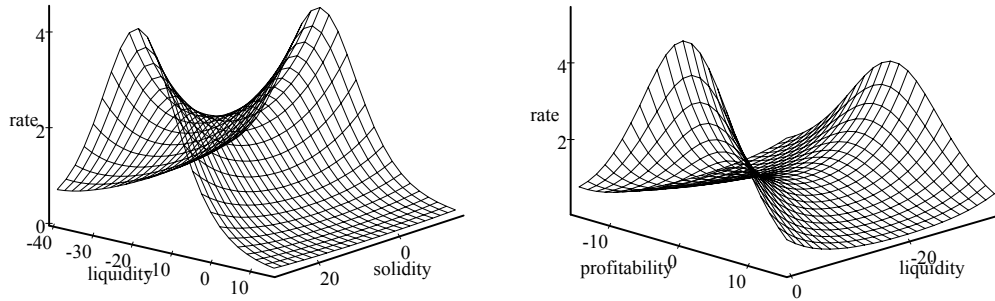
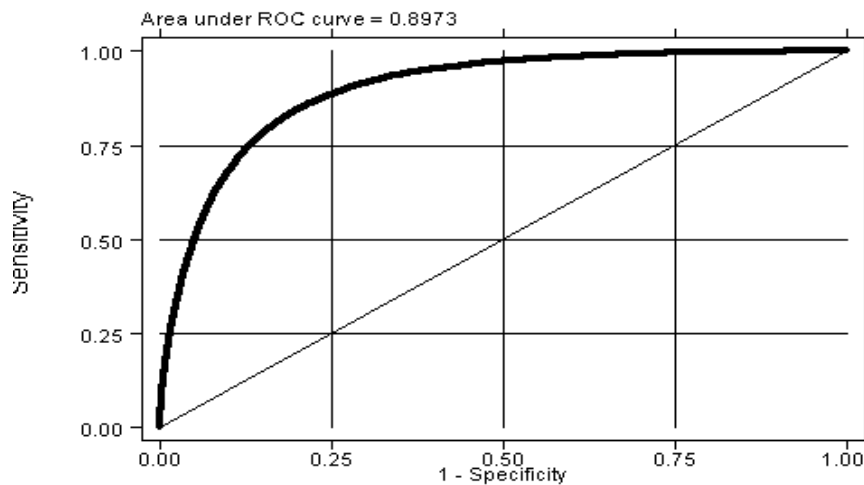


Fig 4: Simulated rates of compensation.

3.6 Predictive ability

If the model is to be used to make binary predictions, a cut-off point for the predicted probability must be determined. The optimal cut-off point will depend on the relative cost of type one and type two errors. Figure 5 gives the menu of trade-offs between correct classifications of the bankrupt cases (sensitivity) and incorrect classifications of the non-bankrupt cases (1-specificity). The area under the curve above the 45 degree line is a common measure of discriminatory power. For the model with no explanatory power the area under the curve will equal 0.5. If predictions are perfect the measure will equal 1.



Figur 5. Discriminatory power

Applied on the estimation sample, the model can correctly identify 83% of both the bankrupt and non bankrupt cases. When the model is estimated over the years 1990-1993 it correctly classifies 82% of both categories in 1996. Only 3% of the observations constitutes bankruptcies. Incorrect classification of 17% of the non-bankrupt cases therefore gives a great number of false predictions. If the cut-off point is determined as to equal the 90% percentile, such that 10% of the sample is classified as bankrupt, one will correctly classify only 63% of the bankrupt cases. However, because the metric subject to prediction is defined by the event "this is the last registered financial statement, and bankruptcy is recorded within 3 years" a firm which is predicted bankrupt in year t is noted as an erroneous prediction if year t does not constitute the last financial statement, even if bankruptcy is recorded within year $t + 3$.

Figure 6 gives a picture of model stability. The model is estimated over the years 1990-1993 and 1990-1996 respectively and the predicted probabilities are plotted. The predicted probabilities in general and the ranking of the firms in particular is not notably affected by this extension of the estimation sample.

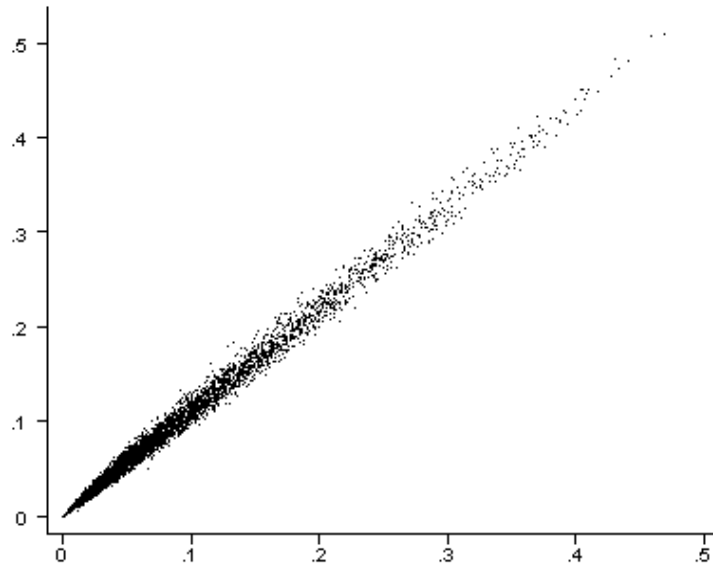


Figure 6. Model stability.

4 Aggregate predictions

In table 3, groups are defined by various intervals of predicted probabilities and mean predicted risk is compared with the observed fractions by year¹². The first order condition for maximizing the log likelihood for the logit model assures that mean predicted probability coincides with proportion of responses in the estimation sample. However, this does not need to hold for every quantile of the predicted probabilities. Table 3 suggests that predicted probabilities fit well with the observed frequencies, not only over the distribution of predicted probabilities, but also over the different years. Despite the differences in the business cycle, there appears to be some degree of stability in the correspondence between mean predicted risk and the observed frequencies over the years 1990-1996.

¹²For the group of firms with a predicted probability of bankruptcy in excess of 20% in 1990, the frequency of bankruptcies was 31 % . For the same group, mean predicted probability were 29%. For the same year; among the firms with a predicted probability less than 1 percent, the frequency of bankruptcies was 0.4 % . Mean predicted probability for this group equaled 0.3 % .

Table 3. Mean predicted risk and observed fractions

Cut-off point	Fraction 1990	Mean predicted 1990	Fraction 1991	Mean predicted 1991
p>.2	.315827	.2939747	.309848	.2976207
.2>p>.1	.186139	.1409671	.170139	.1400238
.1>p>.05	.107498	.0710920	.100804	.0708527
.05>p>.02	.049738	.0322007	.048263	.0322326
.02>p>.01	.021866	.0143543	.019082	.0142045
p<.01	.004291	.0032402	.004241	.0032123
Cut-off point	Fraction 1992	Mean predicted 1992	Fraction 1993	Mean predicted 1993
p>.2	.234043	.2933678	.299776	.2920722
.2>p>.1	.119509	.1389947	.134093	.1390491
.1>p>.05	.065691	.0703857	.071406	.0700618
.05>p>.02	.026517	.0321568	.025918	.0318994
.02>p>.01	.010394	.0142870	.011211	.0142867
p<.01	.002148	.0028966	.002405	.0027531
Cut-off point	Fraction 1994	Mean predicted 1994	Fraction 1995	Mean predicted 1995
p>.2	.271605	.2947980	.188335	.2896208
.2>p>.1	.123089	.1378097	.107410	.1387632
.1>p>.05	.055202	.0698960	.057420	.0703471
.05>p>.02	.024048	.0318248	.027301	.0317281
.02>p>.01	.010627	.0142041	.009455	.0142126
p<.01	.001627	.0026003	.001434	.0026231
Cut-off point	Fraction 1996	Mean predicted 1996		
p>.2	.231141	.2955294		
.2>p>.1	.123243	.1390570		
.1>p>.05	.065801	.0707439		
.05>p>.02	.029142	.0317599		
.02>p>.01	.014248	.0141861		
p<.01	.002258	.0025373		

This result is particularly useful for the effort of predicting expected loss on a portfolio of firms. Figure 9 gives a plot of risk weighted debt 1988-1999. This measure is constructed by multiplying individual probability predictions with the firms debt, and can be interpreted as expected loss related to the event of bankruptcy on the SEBRA portfolio in absence of collateral values to be recovered. Figure 10 gives the reported loan losses of the entire Norwegian banking sector for the same period¹³. Figure 11 gives the fitted regression of loan losses on risk weighted debt with a one year lag. Risk weighted debt is lagged for two reasons. First; the analysis on micro data identified a substantial lag between the last reported statement and the date of bankruptcy for most bankrupt firms. Second; we would expect banks to analyze the statements of year t-1 when determining provisions on loan losses at time t. As the financial statements for year t is available as late as in the middle of year t+1, we need this lag to be able to make predictive statements.

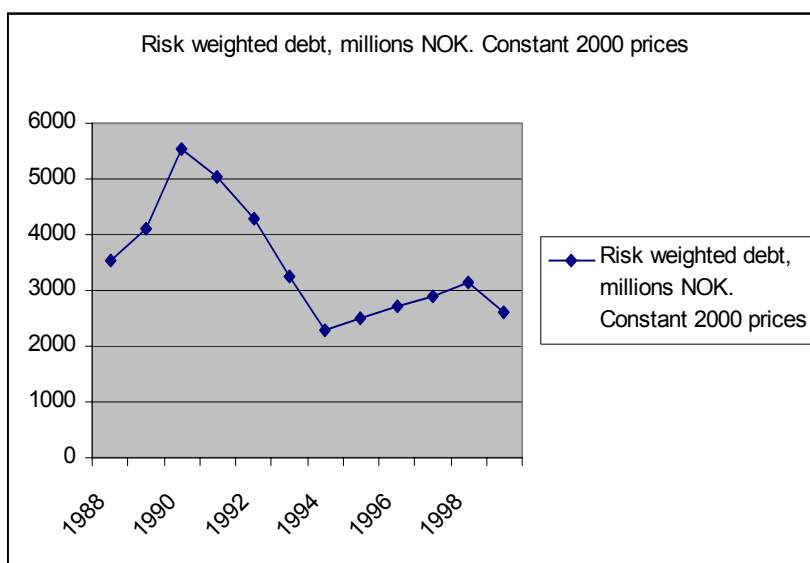


Figure 9

¹³Loan losses of the year and loanloss provisions minus write-backs of previous years loanloss provisions.

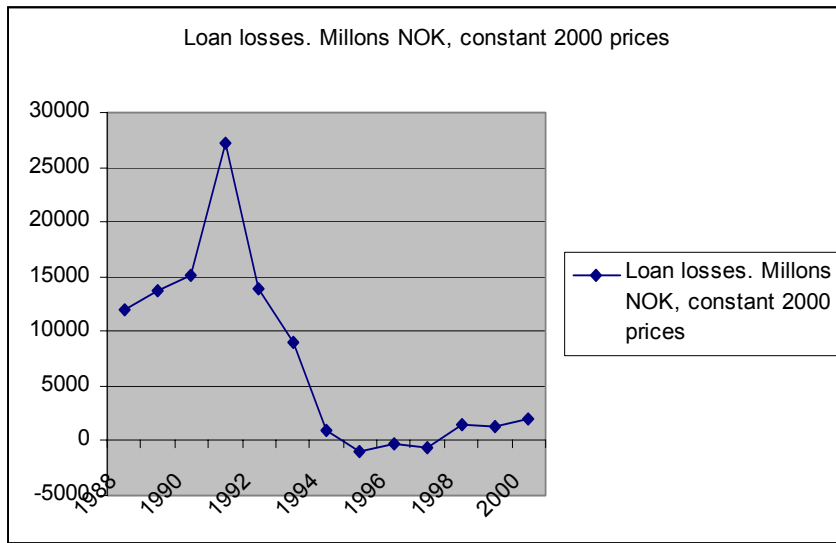


Figure 10

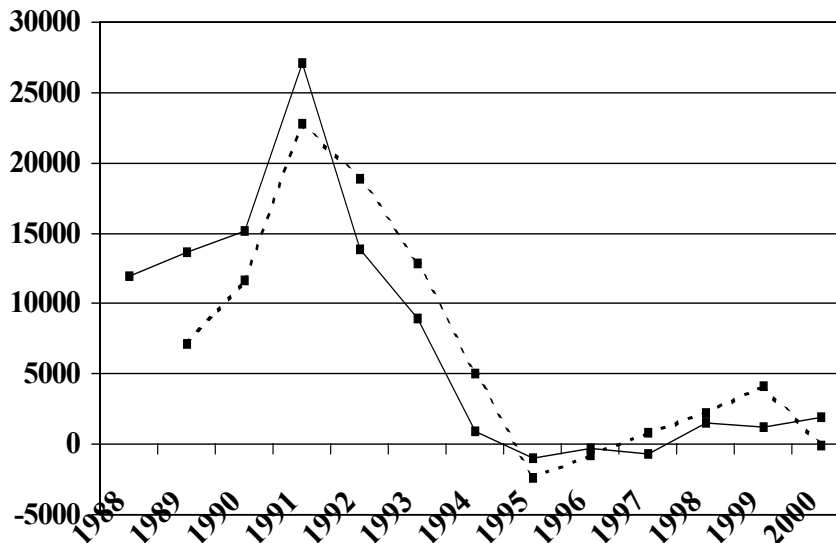


Fig 11. loan losses and fitted loan losses

The R-squared of the regression with a constant term included is 81%. In particular, the regression does fit well with the massive loan losses in 1991. Although one should be careful to draw conclusions on the basis of such a short time-series, this finding is indeed encouraging. Several studies have explored the dependence of aggregate loan losses on

various macro variables¹⁴. However, to the knowledge of the author, not many attempts has been made to explain bank sector loan losses by aggregating micro predictions.

Because there is a tendency of under-prediction at the early years and over-prediction at late years one would expect that inclusion of a macro-variable will show useful in this model. In particular we would like to include a variable that captures the variation in collateral values over the business cycle. In fig 12 loan losses is fitted with lagged risk weighted debt and the change in the real price of housing (the econ index). The macro variable does improve the fit of the model¹⁵.

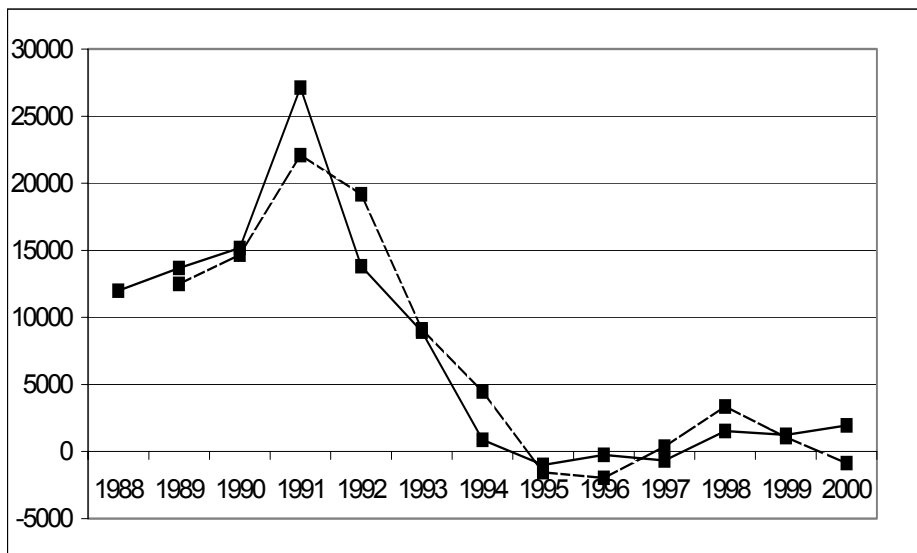


Figure 12. Loan losses fitted with risk weighted debt and change in the econ real price of housing index.

¹⁴See Pesola (2000).

¹⁵The regression output is reproduced in the appendix.

5 Financial distress versus bankruptcy

5.1 The model

The usual specification of the bankruptcy prediction model implicitly assumes that the event of bankruptcy is directly connected to the quality of the financial statements: If the figures are sufficiently bad, we expect the firm to go bankrupt. Yet, as emphasized by the restructuring models, once in a situation of financial distress bankruptcy is only one of several possible outcomes.

Assume that $\beta'x$ measures the financial soundness of the firm, and furthermore that there exist a critical level of this measure such that if this level is exceeded the firm has an insolvency problem. Once in the state of insolvency, assume the claimants on the firm will initiate negotiations, and furthermore that the outcome of these negotiations cannot be predicted by financial statements. Formally: let the event of insolvency be represented by a the binary variable y^o , and assume that the outcome of bankruptcy will occur with a fixed probability conditional on this variable. The critical level of $\beta'x$ is assumed to differ among the firms due to individual characteristics that is not captured by the financial ratios, and thus the error term u is introduced:

$$\begin{aligned}y^* &= \beta'x + u \\y^o &= 1 \text{ if } y^* > 0 \\y^o &= 0 \text{ else}\end{aligned}\tag{25}$$

In this model not only y^* is latent, but also the variable y^o . What is observable is the outcome of bankruptcy represented by the binary variable y , such that:

$$\begin{aligned}P(y = 1|y^o = 1) &= (1 - q) \\P(y = 0|y^o = 0) &= (1 - r)\end{aligned}\tag{26}$$

Thus:

$$\begin{aligned} \text{if } y^* > 0 & \text{ then } y = 1 \text{ with probability } (1 - q) \\ \text{if } y^* \leq 0 & \text{ then } y = 1 \text{ with probability } r \end{aligned}$$

Solving the model for the probability of observing bankruptcy we have:

$$\begin{aligned} P(y = 1|x) &= P(y = 1|y^o = 1) P(y^o = 1) + P(y = 1|y^o = 0) P(y^o = 0) \quad (27) \\ &= r + (1 - q - r) P(u > -\beta'x) \end{aligned}$$

Assuming u is logistically distributed we have;

$$P(y = 1|x) = r + \frac{(1 - q - r)}{1 + \exp(-\beta'x)} \quad (28)$$

By this model, the probability of observing bankruptcy will be constrained to the $[r, (1 - q)]$ interval. Conditional on the model structure, the identification of β allows consistent probabilities of the (unobservable) event of financial distress to be calculated. Independently of how one would interpret the motivated setting, the proposed functional form may be desirable: The transformation of variables suggested in section 3.1 implied that the marginal effect of a single variable should decline as the variable deviates from some critical value at any given level of the probability of bankruptcy. The functional form suggested in this section will have implications as to how the marginal effect of any variable is related to the *over-all* evaluation of the firm: By the structure given in section 3.1, the marginal effect of any variable x_r was given by;

$$\frac{\partial P}{\partial x_r} = P(1 - P)T_r(x_r)(1 - T_r(x_r))\frac{\beta_r}{\delta_r}$$

The marginal effect is dependent on the over-all evaluation of the firm by the term

$P(1 - P)$,which is maximized at $P = 0.5$. By the model of eq(28) , if we impose the restriction of $r = 0$, the marginal effect of variable x_r is

$$\frac{\partial P}{\partial x_i} = P\left(1 - \frac{P}{1 - q}\right)T_r(x_r)(1 - T_r(x_r))\frac{\beta_r}{\delta_r}$$

At any given level of x_r the marginal effect will be greatest when $P^* = 0.5(1 - q)$. If $q > 0$ (still we impose $r + q < 1$), this will imply that marginal effects will be maximized at lower level of P , and thus the *over – all* evaluation of the firm contributes to marginal effects more conservatively as probability estimates grow large.

5.2 Estimation

The structure is analogous to a basic model of misclassification ¹⁶: The log likelihood is given by:

$$\ell((q, r, \beta) = n^{-1} \sum_{i=1}^n \left\{ \begin{array}{l} y_i \ln \left(r + \frac{(1-q-r)}{1+\exp(-\beta'x)} \right) \\ + (1 - y_i) \ln \left(1 - \left(r + \frac{(1-q-r)}{1+\exp(-\beta'x)} \right) \right) \end{array} \right\} \quad (29)$$

For the identification of the vector (q, r, β) , the condition $q + r < 1$ must be imposed, as the estimators based on the maximum likelihood procedure will not be able to distinguish between the parameter values (r, q, β) , and $(1 - r, 1 - q, -\beta)$ ¹⁷. The model is estimated on transformed variables, conditioned on scale and position parameters identified in the previous section Model(A).

¹⁶J.A. Hausman et al (1998)

¹⁷It was found difficult formally to impose this restriction on the standard Stata ML-evaluators. Technically the restriction was imposed by replacing $1 - r - q$ with $(\sqrt{1 - r - q})^2$ computing the log-likelihood equation. In this way the program was kept from evaluating the log likelihood at $r + q > 1$

Estimation results. Model C.

variable	coeficient	s.e.
eka	-1.6161	0.0696
lik	-1.6337	0.0498
tkr	-1.3022	0.0490
lev	0.5111	0.0534
ube	8.2014	0.2823
a1	0.9349	0.0513
a2	1.1107	0.0472
a3	0.9375	0.0475
a4	0.7416	0.0496
a5	0.5723	0.0531
a6	0.3361	0.0588
a7	0.2814	0.0636
a8	0.2058	0.0702
div	-0.9756	0.0752
size	-0.0511	0.0069
taptek	0.5012	0.0438
meanlev	0.9192	0.1921
meaneka	-4.8429	0.2902
sdtkr	2.3266	0.3804
const	-6.7421	0.3244
q	0.5110	0.0286
r	-0.0000783	0.000379

Wald chi(19) = 4492.89

log-L = -29847.179

The probability of bankruptcy given insolvency is estimated to equal 49%, which appears to be reasonable. The probability of bankruptcy given solvency is not significantly different from zero at a 1% level. By this model the marginal effect of any variable x is maximized at $P = 0.245$ for any given level of the same variable x . Figure 13 gives a plot of model A against model C predictions. By the likelihood ratio test, the joint restriction $r = q = 0$ (model A against model C) is rejected.

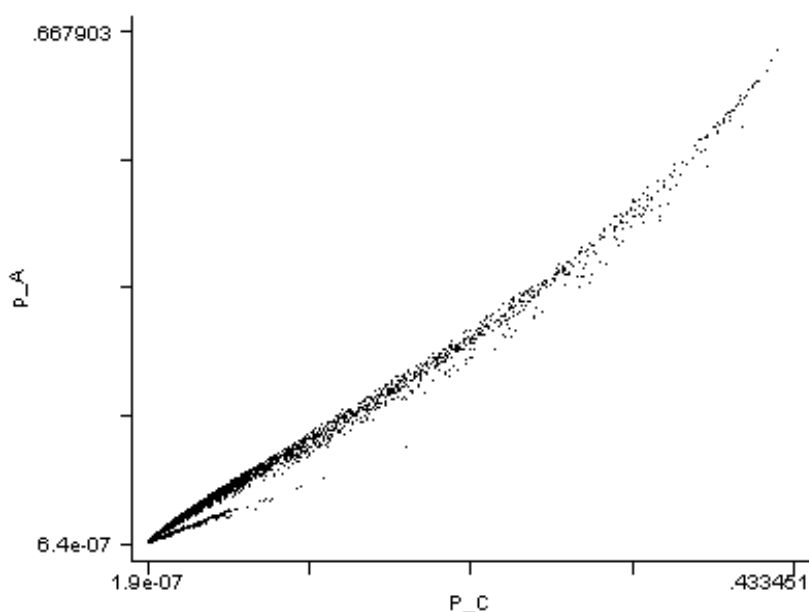


Fig 13 A plot of model A and model C predictions.

6 A suggestion on further research

For most panels of accounts data the number of years will be small and estimating the effects of macro variables will be difficult. Sometimes a longer time-series will be available on aggregated data. In what follows I will explore whether a model estimated on aggregated data can be utilized to identify macro-coefficients for inclusion in the micro-level model.

6.1 Aggregation in the linear regression model

To set focus, I will quickly review aggregation in the case of the linear regression model:

Assume the model:

$$y_{it} = \alpha + x_{it}\beta + z_t\gamma + u_{it} \quad , \quad i = 1, 2, \dots, N \quad , \quad t = 1 \quad (30)$$

With nothing but cross-sectional data available we cannot separate γ from the constant term. Assume however that a time-series is available on both the time specific variable and the mean value of x . By aggregating over individuals in eq(30), we have

$$\bar{y}_t = \alpha + \bar{x}_t\beta + z_t\gamma + \bar{u}_t \quad , \quad t = 1, 2, \dots, T \quad (31)$$

the "between period" estimate of β . Although we are subject to cross-sectional data on the individual characteristics, by including the estimate of γ in eq(30) and adjusting the constant term we can now obtain individual predictions conditioned on future development in the time specific variable, also for the case in which \bar{x} and z are correlated.

6.2 Aggregation in the probit model

Consider the usual pooled probit specification;

$$\begin{aligned} y_{it}^* &= x_{it}\beta + z_t\gamma + u_{it} \quad i = 1, 2, \dots, N \quad , \quad t = 1 \\ y_{it} &= 1 \text{ if } y_{it}^* > 0 \\ y_{it} &= 0 \quad , \quad \text{else} \\ u_{it} &\sim N(0, \sigma^2) \end{aligned} \quad (32)$$

$$\Rightarrow P(y = 1|x_{it}, z_t) = P(y_{it}^* > 0|x_{it}, z_t) \quad (33)$$

$$= P\left(\frac{u_{it}}{\sigma} > -\frac{(\beta x_{it} + z_t \gamma)}{\sigma}\right) \quad (34)$$

$$= \Phi\left(x_{it}\frac{\beta}{\sigma} + z_t\frac{\gamma}{\sigma}\right) \quad (35)$$

If

$$x_{it}\beta \sim N(\mu_t, \sigma_{x\beta}^2) \quad (36)$$

$$\Rightarrow \overline{x_t\beta} \sim N\left(\mu_t, \frac{\sigma_{x\beta}^2}{n}\right) \quad (37)$$

$$\Rightarrow p \lim_{n \rightarrow \infty} (\overline{x_t\beta}) = \mu_t \quad (38)$$

Then there exists parameters $\tilde{\beta}$, and $\tilde{\gamma}$ both identifiable functions of $\frac{\beta}{\sigma}$ and $\frac{\gamma}{\sigma}$ such that for a large N we can write:

$$E_{x_{it}} \Phi\left(x_{it}\frac{\beta}{\sigma} + z_t\frac{\gamma}{\sigma}\right) \simeq \Phi\left(\overline{x_t\beta} + z_t\tilde{\gamma}\right) \quad (39)$$

This is easily seen by rearranging eq(32);

$$y_{it}^* = x_{it}\beta + z_t\gamma + u_{it} = \overline{x_t\beta} + z_t\gamma + (x_{it} - \overline{x_t})\beta + u_{it} \quad (40)$$

By standardizing by the variance of the compound error term, we have:

$$\tilde{y}_{it}^* = \overline{x_t\beta} + z_t\tilde{\gamma} + \omega_{it} \quad (41)$$

Considering that;

$$E(\omega_{it}|z_t, \bar{x}_t) = E\left(\frac{(x_{it} - \bar{x}_t)\beta + u_{it}}{\sigma\sqrt{1 + \sigma_{x\beta}^2}}\right) = \frac{\mu_t - \bar{x}_t\beta}{\sigma\sqrt{1 + \sigma_{x\beta}^2}}$$

, and assuming that N is large, then by eq(38)

$$\bar{x}_t\beta \simeq \mu_t \quad (42)$$

and consequently:

$$\omega_{it} \sim N(0, 1) \quad (43)$$

Considering the identity;

$$\widetilde{y}_{it}^* = \frac{y_{it}^*}{\sigma\sqrt{1 + \sigma_{x\beta}^2}} \quad (44)$$

, obviously $\widetilde{y}_{it}^* > 0$ if and only if $y_{it}^* > 0$, which in turn implies;

$$\begin{aligned} P(y_{it} = 1|z_t, \bar{x}_t) &= P(y_{it}^* > 0|z_t, \bar{x}_t) = P(\widetilde{y}_{it}^* > 0|z_t, \bar{x}_t) \\ &= P\left(\omega_{it} > -\left(\bar{x}_t\widetilde{\beta} + z_t\widetilde{\gamma}\right)\right) = \Phi\left(\bar{x}_t\widetilde{\beta} + z_t\widetilde{\gamma}\right) \end{aligned} \quad (45)$$

Note the difference between the expressions $P(y_{it} = 1|x_{it}, z_t)$ and $P(y_{it} = 1|z_t, \bar{x}_t)$. The former is the probability that individual i will give positive response at time t as a function of both the macro variable and value of the individual characteristic, where as the latter refers to the same probability as a function of the macro variable and the mean value of the characteristic. The latter therefore has the interpretation of being the

expected proportion of positive responses in the population at time t , that is;

$$P(y_{it} = 1|z_t, \mu_t) = E_{x_{it}} (P(y_{it} = 1|x_{it}, z_t)) \simeq \frac{1}{N} \sum_{i=1}^N P(y_{it} = 1|x_{it}, z_t) \quad (46)$$

For \bar{x}_t to be dependent of macro conditions, some form of non-stationary must be assumed for the distribution of $\beta'x_{it}$. In the framework sketched above, the mean value μ is assumed to be time dependent, whereas the variance is assumed to be constant. In other words; changes in over all credit risk is assumed only to shift the mean of the distribution of $\beta'x_{it}$. With reference to elementary finance theory, market portfolios are rarely assumed to have a constant variance, and it is not obvious that portfolios exposed to credit risk will be much different in nature. Different industries will depend in different ways on the macro environment, and thus a macro shock, a policy shock, or an overall change in the business cycle is expected to inflate the variance of $\beta'x_{it}$. Some effort should therefore be devoted to explore the robustness of the model to deviations from the constant variance assumption. It is however not likely that this aspect can be modeled directly, due to the limited time dimension of the panel in hand. Some remedy for this problem should however be brought about if we aggregate by industry. If the aggregation is sufficiently "fine-meshed", heterogeneity in macro dependency over industries could at least partially be considered taken care of.

So far I have sketched a way to start with the estimate of $(\frac{\beta}{\sigma}, \frac{\gamma}{\sigma}, \sigma_{x\beta}^2)$ derived from the micro-level model, and arrive at an aggregate model for mean predicted outcome. By conditioning on future developments of \bar{x} and z the model can be used for evaluating scenarios. Because z is likely to be policy dependent, the model could appear useful for policy analysis. Because the panel is subject to limited time variation, we are however not likely to obtain any sensible estimate of $\frac{\gamma}{\sigma}$. We should therefore seek to find some analogy to the case of the linear regression model sketched in section 6.1. The question is whether we can estimate the model;

$$P_t = \Phi \left(\bar{x}_t \tilde{\beta} + z_t \tilde{\gamma} \right) , t = 1, 2, \dots, T \quad (47)$$

, and identify of the $\frac{\gamma}{\sigma}$ -coefficient by some function of the estimated $\frac{\tilde{\gamma}}{\sigma}$ -coefficient.

The log likelihood for the problem is;

$$l = \sum_{t \in T} \left\{ \bar{y}_t \ln \left[\Phi \left(\bar{x}_t \tilde{\beta} + z_t \tilde{\gamma} \right) \right] + (1 - \bar{y}_t) \ln \left[1 - \Phi \left(\bar{x}_t \tilde{\beta} + z_t \tilde{\gamma} \right) \right] \right\} \cdot n_t \quad (48)$$

where \bar{y}_t and n_t is the fraction of firms going bankrupt and the number¹⁸ of observations at time t respectively.

Alternatively we can turn to minimum chi-square methods. By inverting eq(47), we have;

$$\Phi^{-1}(P_t) = \bar{x}_t \tilde{\beta} + z_t \tilde{\gamma} \quad (49)$$

, and thus we could estimate the equation;

$$\Phi^{-1}(\bar{y}_t) = \bar{x}_t \tilde{\beta} + z_t \tilde{\gamma} + \eta_t \quad (50)$$

as a linear regression model. Expanding $\Phi^{-1}(\bar{y}_t)$ around P_t by a first order Taylor approximation we have:

$$\begin{aligned} \Phi^{-1}(\bar{y}_t) &= \Phi^{-1}(P_t) + (\bar{y}_t - P_t) \frac{\partial \Phi^{-1}(P_t)}{\partial P_t} \\ &= \Phi^{-1}(P_t) + \frac{\bar{y}_t - P_t}{\varphi(P_t)} \end{aligned} \quad (51)$$

$$E(\eta_t) = E\left(\frac{\bar{y}_t - P_t}{\varphi(P_t)}\right) = 0 \quad (52)$$

$$Var(\eta_t) = \frac{P_t(1 - P_t)}{n_t \varphi(P_t)^2} \equiv w_t^2 \quad (53)$$

, where $\varphi(P_t)$ denotes the normal density function. Thus we should apply the method

¹⁸This number will be large, and furthermore vary little over the estimation period. The effect of this weighting should therefore be marginal.

of weighted least squares, substituting \widehat{w}_t for w_t in applying an iterative procedure.

6.3 Identification of the micro model coefficients

After we have obtained a consistent estimate of the $\widetilde{\gamma}$ parameter, we wish to derive the parameter γ for inclusion in the micro-level model. Under the constant variance assumption, this appears to be rather straight forward. Considering that

$$\begin{aligned} \frac{\widetilde{\gamma}}{\sigma} &= \frac{\frac{\gamma}{\sigma}}{\sqrt{1 + \sigma_{x\beta}^2}} \\ \Rightarrow \frac{\gamma}{\sigma} &= \frac{\widetilde{\gamma}}{\sigma} \cdot \sqrt{1 + \sigma_{x\beta}^2} \end{aligned} \tag{54}$$

all that is needed for identification is the estimate of $\sigma_{x\beta}^2$ from the micro-level model.

7 Bibliography

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

Altman, E.I.,Y. Chen and Weston F.(1995):"Financial Distress and Restructuring Models", *Financial Management*,Vol.24,No.2,Summer,1995, p57-75

Beaver, W.(1966):"Financial Ratios as Predictors of Failure", *Journal of Accounting Research*,p.77-111

Bulow, J.I. and J.B. Shoven(1978):"The Bankruptcy Decision" *Bell Journal of Economics*, Autumn,p.4456

Dixit, A.K. and R.S Pindyk(1993): *Investment under Uncertainty*, Princeton University Press, Princeton

Guilkey, D.K. and J.L. Murphy(1993):"Estimating and testing in the random effects probit model", *Journal of Econometrics* ,Vol.59,p.301-317

Hausman, J.A.,J. Abrevaya and Scott-Morton, F.M.(1998):"Misclassification of the dependent variable in a discrete-response setting", *Journal of Econometrics*,Vol.87,p239-269

Laitinen, E.K. and T. Laitinen(2000):"Bankruptcy prediction. Application of the Taylor's expansion in logistic regression", *International Review of Financial Analysis*.Vol.9,p372-349

Maddala, G.S.(1983): *Limited dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge

Maddala, G.S.(1987):"Limited Dependent Variable Models Using Panel Data", *The Journal of Human Resources*..Vol.3,p.307-337

Morris, R.(1997): *Early Warning Indicators of Corporate Failure*, Ashgate Publishing Ltd. Hants

Myers, S.C.(1977):"Determinants of Corporate Borrowing", *Journal of Financial Economics*, November,147-175

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

Pesola,J (2000): "Macro Indicators for Banking Sector Stability in Finland", Bank of Finland , Working paper No. 9.

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

8 Appendix

8.1 Summary statistics

eka				lev			
-----				-----			
Percentiles				Percentiles			
1%	-147.093			1%	0		
5%	-41.30435			5%	0		
10%	-14.41969	Obs	398689	10%	0	Obs	398689
25%	5.41543	Sum of Wgt.	398689	25%	1.474531	Sum of Wgt.	398689
50%	19.94595	Mean	16.24358	50%	11.85065	Mean	21.14765
		Std. Dev.	379.3956			Std. Dev.	53.3773
75%	40.30612	Variance	143941	75%	30.78431	Variance	2849.137
90%	62.90516	Skewness	-319.3263	90%	54.35166	Skewness	116.9907
95%	77.58129	Kurtosis	1105602	95%	70.56429	Kurtosis	22421.64
99%	97.70463			99%	103.2419		
-----				-----			
tlr				ube			
-----				-----			
Percentiles				Percentiles			
1%	-68.88217			1%	0		
5%	-20.96562			5%	0		
10%	-8.870968	Obs	398689	10%	0	Obs	398689
25%	5.683718	Sum of Wgt.	398689	25%	6738328		
50%	6.816901	Mean	6.212922	50%	5.924596	Mean	9.14002
		Std. Dev.	40.48795			Std. Dev.	11.74346
75%	15.05618	Variance	1639.274	75%	12.98763	Variance	137.9088
90%	24.91975	Skewness	-61.31777	90%	22.29508	Skewness	4.87606
95%	32.68698	Kurtosis	13826	95%	30.0995	Kurtosis	149.3876
99%	55.02645			99%	52.61538		
-----				-----			
lik							

Percentiles							
1%	-95700						
5%	-348.2758						
10%	-86.97632	Obs	398689				
25%	-31.43806	Sum of Wgt.	398689				
50%	-15.09607	Mean	-11202.93				
		Std. Dev.	654424.2				
75%	-4.062448	Variance	4.28e+11				
90%	19.04137	Skewness	-83.5831				
95%	85.937	Kurtosis	1.2239.92				
99%	23400						

Table 1. Variables for which transformations are estimated

Variable	Mean	Std.dev	Min	Max
size	2.063386	3.804086	2.05e-10	121.8368
meanlev	.429475	.0926014	.2710669	.6487265
meaneka	.6449002	.0606694	.4502819	.8484569
sdtkr	.2827032	.0454702	.136976	.4127187

Table2:Size and industry characteristics.

Variable	Frequency
a1	.0553339
a2	.0773861
a3	.0793676
a4	.0755025
a5	.0714291
a6	.0663525
a7	.0603352
a8	.052931
div	.335941
taptek	.2892179

Table3.Dummy variables. Fractions.

8.2 Regression output. Loan losses and risk weighted debt.

```
.reg rlossL.rwd
```

Source	SS	df	MS	Number of obs =
Model	717861100	1	717861100	12
Residual	149960726	10	14996072.6	F(1, 10) = 47.87
Total	867821826	11	78892893.2	Prob > F = 0.0000

R-squared = 0.8272
Adj R-squared = 0.8099
Root MSE = 3872.5

rloss	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rwd					
L1	7.727369	1.116864	6.92	0.000	5.238841 10.2159
_cons	-20165.79	4063.666	-4.96	0.001	-29220.2 -11111.38

```
.reg rlossL.rwd econ
```

Source	SS	df	MS	Number of obs =
Model	783582216	2	391791108	12
Residual	84239609.5	9	9359956.61	F(2, 9) = 41.86
Total	867821826	11	78892893.2	Prob > F = 0.0000

R-squared = 0.9029
Adj R-squared = 0.8814
Root MSE = 3059.4

rloss	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rwd					
L1	6.73215	.9589729	7.02	0.000	4.562803 8.901498
econ	-43962.01	16590.6	-2.65	0.026	-81492.55 -6431.474
_cons	-17050.79	3418.908	-4.99	0.001	-24784.89 -9316.68

Regression of risk weighted debt and the change in the real price of housing index on loan losses for the bank sector

KEYWORDS:

Bankruptcy
Logit analysis
Non-linear estimation
Aggregation