Business failure prediction: simple-intuitive models versus statistical models *

Hubert Ooghe[1]
Hubert.Ooghe@UGent.be

Christophe Spaenjers[2]
Christophe.Spaenjers@UGent.be

Pieter Vandermoere[3]
Pieter.Vandermoere@Vlerick.be

(*) The authors thank Graydon Belgium N.V. for providing the data and Sofie Balcaen for her cooperation to this research project in an earlier phase

[1] Professor, Department of Accountancy and Corporate Finance, Ghent University, Belgium; Ernst & Young Chair of Growth Management, Vlerick Leuven Gent Management School, Belgium
[2] Research assistant, Department of Accountancy and Corporate Finance, Ghent University, Belgium
[3] Research assistant, Vlerick Leuven Gent Management School, Belgium

Abstract

We give an overview of the shortcomings of the most frequently used statistical techniques in failure prediction modelling. The statistical procedures that underpin the selection of variables and the determination of coefficients often lead to 'overfitting'. We also see that the 'expected signs' of variables are sometimes neglected and that an underlying theoretical framework mostly does not exist.
Based on the current knowledge of failing firms, we construct a new type of failure prediction models, namely 'simple-intuitive models'. In these models, eight variables are first logit-transformed and then equally weighted. These models are tested on two broad validation samples (1 year prior to failure and 3 years prior to failure) of Belgian companies. The performance results of the best simple-intuitive model are comparable to those of less transparent and more complex statistical models.

"*Complicated procedures do not necessarily provide better results.*"

(Karels & Prakash, 1987, p. 589)

Introduction

From the late 1960s to the present day, failure prediction and financial distress models have been much discussed in the accounting and credit management literature. The topic has developed to a major research domain in corporate finance: since the first failure prediction models of Altman (1968) and Beaver (1967), many studies have been dedicated to the search for the best corporate failure prediction model, based on publicly available data and statistical techniques.

In many countries researchers have attempted to construct an accurate failure prediction model. Altman and Narayanan (1997) mention among others these examples: Ko (Japan, 1982), Fischer (Germany, 1981), Taffler & Tisshaw (UK, 1977), Altman et al. (France, 1974), Knight (Canada, 1979), Fernandez (Spain, 1988), Swanson & Tybout (Argentina, 1988), Gloubos & Grammaticos (Greece, 1988). In Belgium, the first financial distress models were estimated in 1982 by Ooghe and Verbaere. In 1991, Ooghe, Joos & De Vos estimated a second generation of models (Ooghe, Joos & De Bourdeaudhuij, 1995).

All these failure prediction models are largely based on statistical methods. This means that (1) the choice of the variables included in the models is based on a statistical analysis of a certain data set and (2) a coefficient for each variable is estimated by means of a statistical procedure. Balcaen & Ooghe (2004) give an overview of the classical (cross-sectional) statistical methodologies. These include univariate analysis, risk index models, multivariate discriminant analysis and conditional probability models.

Recently, many papers comparing different scoring techniques (applied on the same data set) have been published. Some examples are Bell et al. (1990), Curram & Mingers (1994), Joos, Ooghe & Sierens (1998), Laitinen & Kankaanpää (1999). In addition, some attention has been paid to the comparison of the performance of different types of failure prediction models (Mossman et al., 1998).

If we compare the performance of different failure prediction models based on statistical methods on a same sample, we find that the performance results of most of the statistical methods are quite similar (Ooghe & Balcaen, 2002; Platt & Platt, 1990). Laitinen & Kankaanpää (1999) even argue that the latest applications are as effective in predicting business failure as discriminant analysis was in the late '60s. Since there are no important differences in the predictive abilities of statistical models, it is important to analyze the problems related to their use.

Therefore, in this paper we give an overview of the problems and shortcomings of the most frequently used statistical techniques. Then we investigate what happens if we drop the coefficients in failure prediction models. We make a non-statistical and well-balanced selection of variables, based on expertise of the financial situation of firms, especially of failing ones. We use 'common sense' and, instead of fitting the model to a certain data set, we use the 'expected signs'[4] of the ratios. Finally, we compute the performance of these simple-intuitive models on a data set of Belgian financial statements. The error rates (type I, type II and unweighted error rate) and the Gini-coefficients are compared and the predictive ability of the best simple-intuitive model is compared with the performance level of other, statistical models.

Our research hypothesis is that these 'naïve' or intuitive business failure prediction models perform as good as the more sophisticated statistical models. This, combined with the facts that they are easy to understand and to compute and that they are intuitively more correct, would make them superior to the existing statistical models in predicting business failure.

This paper is structured as follows. Section 1 describes the problems associated with the use of statistical techniques in failure prediction. In section 2 the general build-up of simple-intuitive models is discussed. The performance measures are explained in section 3, while section 4 discusses the population and samples. Section 5 compares the performance results of the different simple-intuitive models. Also in this section the performance of the best simple-intuitive model is compared to that of some statistical models.

1. Statistical models and their related problems

1.1 Statistical models: multiple (linear) discriminant analysis and logit analysis

The two most frequently used statistical techniques in business failure prediction are multiple linear discriminant analysis (MLDA) and logit analysis (LA).

Most failure prediction models use multiple discriminant analysis (in one form or another) to classify observations (annual accounts) in two a priori defined and mutually exclusive groups (failing or non-failing). This happens based on a combination of independent variables (financial ratios).

[4] We expect that a certain ratio has a positive (negative) sign if it is generally assumed to be positively (negatively) correlated with the financial health of a company.

An MLDA model consists of a linear combination of variables. The values of these variables are combined into one discriminant score. This score gives an indication of the financial health of the firm. The discriminant score is used to differentiate between firms that are expected to fail and those expected not to fail in the foreseeable future. So, a certain cut-off point has to be set.

The general linear discriminant function is the following:

$$D = d_0 + d_1 V_1 + d_2 V_2 + \ldots + d_m V_m \tag{1}$$

with D = discriminant score between - ∞ and + ∞;

$V_1 \ldots V_m$ = independent variables of the model;

$d_0 \ldots d_m$ = linear discriminant coefficients.

The pioneering study in this respect is Altman (1968): Altman's Z-score is a well-known example of an MLDA model.

LA is one type of the so-called 'conditional probability models'. These models allow estimating the probability of company failure conditional on a range of firm characteristics. In LA, a non-linear maximum likelihood estimation procedure is used to obtain the estimates of the parameters of the logit model. The LA model combines several characteristics into one probability score:

$$L = \frac{1}{1 + e^{-(b_0 + b_1 V_1 + b_2 V_2 + \ldots + b_m V_m)}} \tag{2}$$

with L = logit score between 0 and 1;

$V_1 \ldots V_m$ = independent variables of the model;

$b_0 \ldots b_m$ = regression coefficients.

Pioneering studies are Martin (1977) and Ohlson (1980).

1.2 Problems and shortcomings of statistical models [5]

Firstly, the use of these statistical techniques is valid only under very restrictive assumptions. Karels & Prakash (1987) are (among others) focusing on the effect of violation of these assumptions.

Secondly, one can question the role of the estimation of coefficients in business failure models. In a recent study of Ooghe & Balcaen (2002) it becomes clear that the re-estimation of the coefficients (on another data set) can lead to very different results. This is related to the well-known problem of overfitting, i.e. "optimising fit to the presented problem, which is merely a single point sample from the space of possible (future) problems" (Hand, 2004). Moreover, if we analyse the signs of the coefficients of failure prediction models, it appears that these do not always correspond to what generally may be expected (examples include the studies of Bilderbeek, 1979; Zavgren, 1985; Gloubos & Grammatikos, 1988; Keasy & McGuinness, 1990; Doumpos & Zopoudinis, 1999).

[5] This paragraph is partly based on Balcaen & Ooghe (2004).

Therefore, we conclude that the estimation of both the sign and the (absolute) value of the coefficients for each of the variables in a failure prediction models is sometimes nothing more than a pure statistical procedure.

Finally, we also question the way in which the variables included in statistical business failure models are selected. Very often researchers start by forming a very wide range of possible variables and then reduce this range to a limited number of variables using one or more statistical techniques. This results in models with sample-specific variables that fit the data set that is used, but that are not suitable using other data sets. The selection of the variables is not based on a general accepted framework or theory about which variables really indicate (financial) problems within companies.

We conclude that using statistical methods implies problems that cannot be ignored. Therefore, we turn to a new sort of models that do not explicitly use model coefficients. All ratios are thus equally weighted in these new "simple-intuitive models".

2. Simple-intuitive models and their variables

This study aims to validate failure models that consist of a number of financial ratios that represent different aspects of the financial situation of a company. As we do not estimate model coefficients, we will have to combine the values of these ratios for each firm j into one model score S_j. A high score S_j indicates that the company is in good shape and is less likely to fail, while a low model score S_j is a warning sign for companies facing financial difficulties and having a high failure probability.

2.1 List of ratios

In order to construct a range of simple failure prediction models, we start from a list of 18 ratios that represent the various aspects of financial health: added value, profitability, solvency and liquidity. These ratios were chosen after a careful examination of the knowledge about financial indicators of high-risk firms. Hence the list includes ratios that are frequently used by financial analysts, that are often focused on in the literature on financial statement analysis or that have proven to be relevant in earlier research on business failure models. Since we want the models to be 'simple', the following list only includes ratios that are understandable and not difficult to compute.

Most ratios are positively related to financial health. Although, for some ratios a high value x_{ij} for firm j indicates a bad financial situation. These ratios thus have a negative 'expected sign'.

Table 1 lists the ratios used in this study. Appendix 1 describes how these ratios are calculated based on the annual account sections in the Belgian financial statements.

Table 1: List of ratios

Ratio		Expected sign
1.	Gross added value / personnel employed (in 000 EUR)	+
2.	Personnel charges / personnel employed (in 000 EUR)	-
3.	Gross added value / value of production	+
4.	Gross added value / personnel charges	+
5.	Financial leverage (= net return on total assets before taxes – average interest rate of debt)	+
6.	Net return on total assets before taxes (= earnings before interests and taxes (EBIT) / total assets)	+
7.	Net return on equity after taxes (= net profit after taxes / shareholder's equity)	+
8.	Gross return on equity after taxes (= cash flow after taxes / shareholder's equity)	+
9.	Self-financing level (= (accumulated profit/losses and retained earnings) / total assets)	+
10.	General level of financial independence (= shareholder's equity / (shareholder's equity + liabilities)	+
11.	Debts guaranteed / total debt	+
12.	Short term financial debt level (= short term financial debt to credit institutions / short term debt)	-
13.	Cash flow after taxes / liabilities	+
14.	Free cash flow (=cash flow after taxes – investments in fixed assets) / financial debt	+
15.	Overdue taxes and social security charges / taxes, remuneration and social security debt	-
16.	Current ratio (= current assets / short term liabilities)	+
17.	(Cash + short-term investments) / total assets	+
18.	(Cash + short-term investments – short-term financial debt) / current assets	+

2.2 Logit transformation

It is important to mention that, when calculating the model score S_j for firm j, we can not simply add up all values R_{ij} of the ratios included in the model, and this for two reasons.

Firstly the adding up of 'positive' and 'negative' ratios would result in a meaningless model score S_j. As we want a high model score to indicate good financial health, we have to take account of the sign that corresponds to each ratio: we use a plus sign for each 'positive' ratio and a minus sign for each 'negative' ratio.

Secondly, as we do not use coefficients and thus all ratios are attributed the same weight, it is clear that all ratios have to fit the same scale. Otherwise some ratios would contribute much more to the model score than others. Consequently, all ratios are rescaled by means of a logit transformation:

$$L_i = \frac{1}{(1+e^{-R_i})} \tag{3}$$

with L_i = logit value of ratio R_i;

R_i = ratio i with its positive or negative sign depending on the presumed positive or negative relationship with the financial situation.

By doing so, all ratios take values between 0 and 1[6]. Some examples of the logit transformation are shown in table 2 below.

Table 2: Logit transformation of ratios

Ratio R_i	Logit value L_i
+ 10	1.0000
+ 5	0.9933
+ 1	0.7311
+ 0.5	0.6225
0	0.5000
-0.5	0.3775
- 1	0.2689
- 5	0.0067
- 10	0.0000

Each firm j in the sample is attributed a logit value L_{ij} for each ratio R_{ij} and we calculate the model score S_j using the logit values instead of the original values of the ratios[7].

It is important to mention that for some annual accounts one or more ratios R_i cannot be calculated because of zero values in the denominator (for example, ratios 1, 2 and 4 for firms without personnel). Also, we have to watch out for negative values in the denominator, which can finally result in positive values for the ratio, due to negative values in the numerator (for example, ratios 7 and 8 for firms with a negative shareholder's value due to losses). In these special cases, where the denominator is equal to or less than 0, the numerator determines the logit value of the ratio:

- If the numerator of the ratio > 0, then $L_{ij} = 1$;
- If the numerator of the ratio = 0, then $L_{ij} = 0.5$;
- If the numerator of the ratio < 0, then $L_{ij} = 0$.

In this way, we can use in this study as many available annual accounts as possible. This procedure also enlarges the applicability of the model.

There are, however, some other general rules that need to be taken into account. For ratio 3 (gross added value / value of production), the value is considered as invalid if the numerator is equal to the denominator. Ratio 5 (financial leverage) is considered invalid if one of the two composing factors has a denominator equal to 0. Firms with invalid values are not included in the samples (cf. infra).

[6] It is clear that even after the logit transformation not all ratios have the same distribution or even cover the same range. There are, for example, ratios that can not be negative, while other can go below zero. This problem can be solved in future research, but in this paper the most general simple-intuitive models are explored.

[7] When calculating the logit values, we may have to make a correction for some ratios before we can make the logit transformation. As ratio values R_{ij} that are larger than +10 or smaller than –10 are transformed into logit values L_{ij} of respectively 1 or 0, we want to make sure that the ratios R_i mostly have values between –5 and +5. Therefore, ratio 1 and ratio 2 are transformed: their values are divided by the average of the year in which the account is published.

2.3 Model score of simple-intuitive models

As we want the total model score to have a value between 0 and 1, we divide the sum of the logit values by the number of ratios used in the model:

$$S_j = \frac{\sum_{i=1}^{n} L_{ij}}{n} \tag{4}$$

with S_j = model score of firm j;

L_{ij} = logit value of ratio i for firm j;

n = number of ratios used.

3. Performance measures[8]

The performance of a classification model indicates how well the model performs and is called 'goodness-of-fit' in the econometric literature. In this study, two different kinds of performance measures will be used: (1) the type I, type II and unweighted error rates, which are based on a 'classification rule' and (2) the Gini-coefficient, which is based on the 'inequality principle' (Joos, Ooghe and Sierens, 1998).

3.1 Measures based on a classification rule

In our model, a high score indicates a healthy financial situation, while a low score indicates a bad financial situation. Thus a firm has a high failure probability and therefore will be classified into the failing group or 'group 0' if its score S_j is lower than a certain cut-off point S^*. Conversely, a company will be classified into the non-failing group or 'group 1' if its score S_j is higher than the cut-off point S^*.

Two types of misclassifications can be made:

- A type I error represents a 'credit risk': a failing firm is classified as a non-failing one (in 'group 1');
- A type II error represents a 'commercial risk': a non-failing firm is classified as a failing one (in 'group 0').

In this respect, the optimal threshold or 'optimal cut-off point' S^* of a failure prediction model can be calculated as the point at which the unweighted average of both types of errors - the 'unweighted error rate' (UER) - is minimized. This optimal cut-off point S^* corresponds to the score for which the greatest difference ($D_{\text{non-failing, failing}}$) between the cumulative distributions of the scores of non-failing firms ($F_{\text{non-failing}}$) and those of the failing firms (F_{failing}) exists.

[8] This section is largely based on Ooghe & Balcaen (2002).

In this study, we use the UER because this is the most objective performance measure. The allocation of weights to the different types of errors is subjective and depends on the degree of risk aversion of the risk analyst. Furthermore, we do not want to take into account the population proportions because of the unbalanced proportion of failing and non-failing companies[9]. The over-representation of non-failing companies would lead to a focus on the minimization of type II error rates, and hence, to cut-off points that are too low and a decision process that is too tolerant.

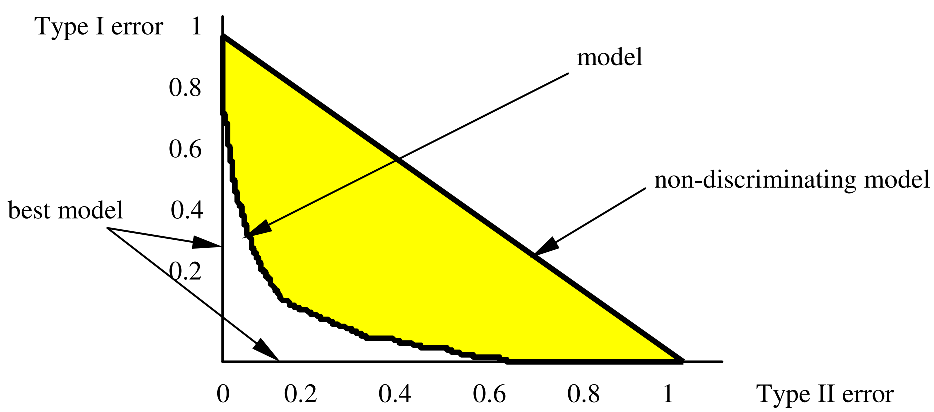
3.2 Measures based on the inequality principle

The performance of a model can also be demonstrated graphically with the construction of a trade-off function (Figure 1). Here, the cumulative frequency distributions from the lowest to the highest scores for 'non-failing' and 'failing' firms are located in a co-ordinate system, with the type II error ($=F_{\text{non-failing}}(y)$) on the X-axis and the type I error ($= 1-F_{\text{failing}}(y)$) on the Y-axis (Steele, 1995),

with $F_{failing}(y)$ = cumulative distribution function of the scores of the failing firms;

$F_{non\text{-}failing}(y)$ = cumulative distribution function of the scores of the non-failing firms.

Figure 1: Trade-off function of a model



It is clear that the best-performing (i.e., most discriminating) model has a trade-off function that coincides with the axes. By contrast, the non-discriminating model, which cannot distinguish between non-failing and failing firms, has a linear descending trade-off function from 100% type I error to 100% type II error. Comparing the location of the trade-off function of a failure prediction model with the location of the most discriminating and the non-discriminating models gives a clear indication of the performance of the model: a model is more accurate if its curve is located closer to the axes.

[9] This also means that the UER does not indicate the real percentage of the firms that is classified falsely by the model.

The Gini-coefficient of a model is an aggregated performance measure that reflects the difference between the trade-off function of the model and the trade-off function of the non-discriminating model. In a normal situation, this coefficient lies between zero and one – it is equal to the proportion of the area between the model and the non-discriminating model (i.e., the grey area in Figure 1) to the area between the non-discriminating and the best model (i.e., the triangle with the axes as sides). As a result, a higher Gini-coefficient corresponds to a curve that is situated closer to the axes, and hence, to a better performing model. An empirical approximation of the Gini-coefficient is shown in the formula below (Joos, Ooghe and Sierens, 1998):

$$\hat{GINI} = \frac{\frac{x_{\max} y_{\max}}{2} - \sum_{i=1}^{n} (x_i - x_{i-1}) \frac{y_{i-1} + y_i}{2}}{\frac{x_{\max} y_{\max}}{2}} \qquad (5)$$

$$= 1 - \sum_{i=1}^{n} (x_i - x_{i-1})(y_{i-1} + y_i)$$

with x_i = type II error rate with threshold i;

y_i = type I error rate with threshold i;

x_{max} = maximum type II error rate, i.e., 100%;

y_{max} = maximum type I error rate, i.e., 100%.

The Gini-coefficient of a model corresponds to the proportion of the area between the cumulative distributions for all scores of non-failing and failing firms to the maximum area of the best model. It is thus based on the differences for all possible scores and not only for the optimal cut-off score, although for most models the unweighted error rate and the Gini-coefficient give similar performance results.

4. Population and samples

4.1. Population and samples of failing and non-failing companies

As we wanted to start from an extensive population of companies over a long time period, Graydon N.V. delivered the VAT numbers of all companies that have closed at least one annual account in the period from January 1, 1990 to December 31, 2001[10]. Table 3 gives the total number of companies in the different industry populations in this study and indicates the NACE-BEL industry codes.

[10] It should be mentioned that some industries were excluded form the analysis: 'public administration and defense', 'education' and 'extra-territorial organizations and bodies'. These are industries with special characteristics where financial distress only rarely leads to bankruptcy.

Table 3: Total population of companies in the different industries 1990-2001

	Industry	NACE-BEL codes	Number of companies
1	Agriculture	01 + 02 + 05	6,007
2	Utilities (energy and water supply)	10 + 11 + 12 + 40 + 41	1,267
3	Metal industry	13 + 27 + 28 + 29 + 30 + 31 + 32 + 33 + 34 + 35 + 371	9,783
4	Food industry	15 + 16	4,156
5	Chemicals	143 + 144 + 145 + 23 + 24 + 25 + 372	2,453
6	Textiles and apparel	17 + 18 + 19	2,663
7	Timber and furniture industry	20 + 361 + 3662	2,719
8	Paper and printing	21 + 22	4,792
9	Other industries	362 + 363 + 364 + 365 + 3661 + 3663	793
10	Construction	141 + 142 + 26 + 45	30,503
11	Wholesale	51	41,458
12	Retail	50 + 52	51,597
13	Hotel, restaurant and catering	55	17,478
14	Transportation	60 + 61 + 62 + 63	20,534
15	Real estate	70	27,864
16	Business services	64 + 67 + 71 + 72 + 73 + 74 – 74151 + 90	59,994
17	Personal services	92 + 93 + 95	10,186
18	Financial services	65 – 65234 + 66 + 67	16,580
19	Health and public services	85	8,515
21	Portfolio companies and management activities of holdings	65234 + 74151	3,634
	TOTAL		**322,976**

Corresponding to the evolution of judicial situations[11], each industry population of companies is divided into the following three groups: a group of **failing** firms, a group of **non-failing** firms and a group of doubt-causing firms. The last group can be again split up into two: a group of **doubt-causing failing** and a group of **doubt-causing non-failing** firms.

A firm is included in the **failing** group if the firm is characterized by one or more of the following judicial situations in the period between January 1, 1990 and December 31, 2001:

- Request for a judicial composition (only used before 1997) (if not returned to a normal condition);
- Official approval of a judicial composition (only used before 1997);
- Temporary postponement of payments (if not returned to a normal condition);
- Final postponement of payments;
- End of the postponement of payments (if not returned to a normal condition);
- Bankruptcy (if not returned to a normal condition);
- Closure of a bankruptcy;
- Other solvency problems (if not returned to a normal condition).

[11] The information concerning the judicial situations of the companies is also obtained from Graydon N.V.

A firm is included in the **doubt-causing failing** group if it is not in the failing group and characterized by one or more of the following judicial situations in the period between January 1, 1990 and December 31, 2001:

- Return to a normal condition;
- Request for a judicial composition (if returned to a normal condition);
- Temporary postponement of payments (if returned to a normal condition);
- End of the postponement of payments (if returned to a normal condition);
- Bankruptcy (if returned to a normal condition);
- Recall of the bankruptcy;
- Other solvency problems (if returned to a normal condition).

A firm is included in the **doubt-causing non-failing** group if it is not in the failing or in the doubt-causing failing group and is characterized by one or more of the following judicial situations in the period between January 1, 1990 and December 31, 2001:

- Termination of activity;
- Voluntary liquidation and dissolution;
- Merger with another company to form a third one;
- Absorption by another company;
- Legal dissolution;
- Closing of a liquidation;
- Scission into several companies;
- Dissolution by legal ending;
- Dissolution without liquidation;
- No apparent activity.

A firm is considered to be **non-failing** if it is not characterized by one or more of the judicial situations mentioned above in the period between January 1, 1990 and December 31, 2001.

The total population of failing and non-failing firms consist of 292.003 non-failing companies (of which 252.496 "pure" non-failing companies and 39.507 doubt-causing non-failing companies) and 30.973 failing companies (of which 30.664 "pure" failing companies and 309 doubt-causing failing companies). An overview of the numbers of companies and accounts in the population and in each sample can be found in appendix 2.

Because we want to estimate different kinds of failure prediction models that need validation, two different samples are required: a sample for estimation[12] and one for validation. Before sampling the data set, we

[12] This estimation sample serves primarily as a means to construct the statistical models to which our 'simple-intuitive models' will be compared. It also offers help in constructing our simple-intuitive models (cf. infra).

randomly reduce the size of the total population of non-failing firms to one third of its original size, since a too large database would be practically unmanageable. Secondly, as we will see later, we only use annual accounts from the period 1990-1999. Considering this explicit timeframe, we move all companies that failed between January 1, 2000 and December 31, 2001 to the sample of doubt-causing non-failing companies for the 1990-1999 period. So, these cases are both in the sample of failing companies (period 2000-2001) and the sample of non-failing companies (1990-1999). Afterwards, the total population of failing and non-failing firms is randomly split into two: one sample for estimation and one sample for validation. Finally, all the doubt-causing cases are skipped from the failing and non-failing estimation samples. By doing so, we can build a model based on data that are as "pure" as possible. The validation sample on the other hand still contains the doubt-causing cases, since this sample needs to be as broad as possible. Table 4 shows the total numbers of companies in the two samples.

Table 4: Total numbers of companies in the failing and the non-failing samples (1990-2001)

	Estimation sample	Validation sample
Number of failing companies	15,348	15,543
Number of non-failing companies	42,226	49,948

4.2 Samples of failing and non-failing annual accounts

After having selected a number of companies, we also have to determine which annual accounts we are going to use for the estimation and validation of the models.

4.2.1 Sample of failing annual accounts

As it is our aim to estimate and validate the models 1 and 3 years prior to failure, it is clear that we need the annual accounts 1 and 3 years prior to failure for each company in the failing sample. The result is two samples of failing annual accounts: a sample of **failing annual accounts 1 year prior to failure** (or '1 ypf') and a sample of **failing annual accounts 3 years prior to failure** (or '3 ypf'). Here, we apply a specific definition of the annual accounts 1 and 3 years prior to failure, because not all companies deposit their annual accounts on December 31:

> *Account one year prior to failure*: account with the closing date falling within the period [date of failure, date of failure – 365 days]
> *Account three years prior to failure*: accounts with the closing date falling within the period [date of failure – (2 * 365 days), date of failure – (3 * 365 days)]

It is important to set a specific timeframe: the annual accounts 1 and 3 years prior to failure have to refer to the same period. Hereby performance measures can be compared. In this study, we only use annual accounts from the period 1990-1999. More recent data concerning the judicial situation were not available at the time of sampling[13].

As we want the failing annual accounts 1 and 3 years prior to failure to refer to the time frame 1990-1999, we first exclude the companies with a failure date after December 31, 1999 from the sample of 15,348 (estimation) and 15,543 (validation) failing annual accounts 1 year prior to failure. This reduces the number of failing annual accounts 1 year prior to failure to 11,492, respectively 11,528. On the other hand, the companies with a failure date before December 31, 1992 are excluded from the sample of failing annual accounts 3 years prior to failure, which reduces the number of failing annual accounts 3 years prior to failure to 14,151 for the estimation sample and 14,363 for the validation sample.

Finally, we eliminate all cases for which the annual account 1 or 3 years prior to failure has not been deposited at the National Bank of Belgium and therefore are not available. This significantly reduces the original number of failing annual accounts in the sample 1 year prior to failure, as many failing companies cease to pay attention to financial reporting when they are close to failure. The original number of failing annual accounts 3 years prior to failure is also reduced. Tables 5 and 6 give the total numbers of annual accounts 1 and 3 years prior to failure.

Table 5: Total numbers of annual accounts (1990-1999) in the estimation sample 1 ypf and 3 ypf

	Total number of annual accounts	Number of available annual accounts
Failing sample 1 ypf (1990-1999)	11,492	2,591
Failing sample 3 ypf (1993-2001)	14,151	10,522

Table 6: Total numbers of annual accounts (1990-1999) in the validation sample 1 ypf and 3 ypf

	Total number of annual accounts	Number of available annual accounts
Failing sample 1 ypf (1990-1999)	11,528	2,676
Failing sample 3 ypf (1993-2001)	14,363	10,624

4.2.2. Sample of non-failing annual accounts

When selecting the two samples of non-failing annual accounts - one for the estimation (1 and 3 year prior to failure) and one for the validation (1 and 3 years prior to failure) - we have to make sure that these annual accounts refer to the same time frame as the failing annual accounts: 1990-1999. First, we exclude all (non-

[13] To use the annual accounts of 2000 and 2001, we would have to know the judicial situation of the company in 2002 and 2003, i.e. three years later. This information was not available at the time.

failing) companies that were started up between January 1, 2000 and December 31, 2001, since they can not provide any annual accounts for the period 1990-1999. Secondly, the samples of non-failing companies are randomly divided into 10 equal groups and for each group of companies, the annual accounts of one specific year in the period 1990−1999 are taken[14]. By doing this, for some companies, the annual accounts could not be provided for the simple reason that they had not been established yet in the chosen year. Although this leads to some loss of data, we are convinced that this is the best method, since we want to avoid young established companies becoming overrepresented in our sample.

Here, we also eliminate all non-failing cases for which the selected annual accounts have not been deposited at the National Bank of Belgium and hence are not available in the database of Graydon N.V. This finally results in a reduced number of non-failing annual accounts in the estimation and validation samples. Table 7 shows the total number of annual accounts, before and after the elimination of non-available annual accounts, in the samples of non-failing annual accounts 1 and 3 years prior to failure.

Table 7: Total numbers of annual accounts (1990-1999) in the samples 1 ypf and 3 ypf

	Total number of annual accounts 1990-2001	Total number of annual accounts 1990-1999	Number of available annual accounts 1990-1999
Non-failing estimation sample	42,226	39,609	27,898
Non-failing validation sample	49,948	47,310	31,946

4.2.3. Final validation samples of annual accounts

Some accounts have invalid results for ratio 5 – these accounts are therefore excluded. This results in one single validation sample that will be used to validate all models. Table 8 reports the number of annual accounts in this validation sample.

Table 8: Total numbers of annual accounts that are used in validation samples 1 ypf and 3 ypf

	Non-failing annual accounts	Failing annual accounts
Sample 1 ypf	31,422	2,656
Sample 3 ypf	31,422	10,510

5. Construction of the models and performance results

[14] There are many possible annual accounts that we can use for each non-failing company in the two non-failing samples.

When constructing the different simple-intuitive models, we decided to exclude ratios 1, 2 and 3. Ratios 1 and 2 were excluded because the denominator of about 60% of the total population is not known. This is caused by the fact that these companies have not filled in the notes item "number of personnel employed", especially before 1996. The denominator (value of production) of ratio 3 cannot be calculated for most of the small firms with an abbreviated form of annual accounts (about one third of the total population).

5.1 Univariate analysis and correlation analysis of the ratios

Before we construct the simple-intuitive models, we list the 15 remaining ratios according to their discriminating power (based on the estimation samples after logit transformation) in appendix 3. This is useful as means of guidance for the construction of the models.

Based on appendix 3 the following conclusions can be drawn:

- 1 year prior to failure ratios haves more discriminating power than 3 years prior to failure ratios;
- the most discriminating ratios in descending order for 1 and 3 years prior to failure are: the general level of financial independence (L_{10}), the cashflow coverage of debt (L_{13}), the net return on equity after taxes (L_7), the self-financing level (L_9) and the gross return on equity after taxes (L_8).

A correlation matrix for the ratios 4 to 18 both 1 year and 3 years prior to failure (based on the total estimation samples) can be found in appendix 4. The intercorrelations between the 15 remaining ratios are rather low. There is a restricted number of intercorrelations higher than 0.60. This means that the 15 selected ratios measure several aspects of the financial situation.

The following ratios are intercorrelated, as can be expected:

- the net return on equity after taxes (ratio 7), the net return on total assets (ratio 6) and the financial leverage (ratio 5), which is the connection between ratio 6 and ratio 7;
- the net return on equity after taxes (ratio 7) and the gross return on equity after taxes (ratio 8), because of the interdependence between both;
- the gross return on equity (ratio 8), the general level of independence (ratio 10) and the cashflow coverage of debt (ratio 13): the higher the gross return on equity and the higher the general level of independence, the higher the cashflow coverage of debt.

5.2 Construction of multivariate simple-intuitive models

As we want to build equilibrated models, we have to select different groups of ratios. An important issue is the number of ratios to be included in the models. When constructing the different models, we want the models to be multi-dimensional and thus include all different types of ratios. This means that we want every possible model to cover added value, profitability, solvency and liquidity.

Also, we want the models to be stable and well balanced. Therefore, we will construct models that include 8 ratios. Due to the construction of our sample we cannot validate the time-stability, but we argue that models with only 2 to 4 ratios will not have a solidity comparable to a model with 8 ratios. We expect models with fewer ratios to result in larger differences when calculating model scores of one company in consecutive years.
On the other hand, in our view, adding even more ratios does not add value to the models. In appendix 5 we show the results of the models with the 1, 2, 3, … 12 most discriminating ratios. This clearly shows that adding ratios does not always increase the performance results of the models.

We build multi-dimensional models based on a combination of 8 ratios with respect to the 4 aspects of the financial situation. Table 9 gives an overview of the 8 different ratios included in the different models tested in this study. Multiple combinations are possible. As ratios 1, 2 and 3 were excluded (cf. supra), ratio 4 is the only remaining added value ratio and therefore was maintained in all models.

Table 9: Different simple-intuitive models (SIM) and their composition (0 not included, + included with a positive sign, - included with a negative sign)

		Models																			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	5	+	+	+	+	+	0	+	0	+	0	0	0	0	0	0	+	+	+	+	+
	6	+	+	0	+	+	0	0	+	0	+	+	+	0	+	+	0	0	0	+	+
	7	0	+	0	0	0	+	+	+	+	+	0	+	0	0	0	0	+	+	0	+
	8	+	0	+	+	+	+	+	0	0	+	+	0	+	+	+	+	+	0	0	0
	9	+	+	+	0	+	0	+	0	0	0	+	+	+	0	+	+	+	+	0	0
	10	+	+	+	+	+	+	0	0	0	0	+	+	+	+	+	+	0	0	0	0
Ratios	***11***	0	0	0	0	0	0	0	+	0	+	0	0	0	0	0	0	0	+	+	+
	12	0	0	0	0	0	-	0	-	-	0	-	-	-	-	-	-	-	0	0	0
	13	+	+	+	+	+	+	0	0	0	0	+	+	+	+	+	+	0	0	0	0
	14	0	0	0	0	0	0	+	+	+	+	0	0	0	0	0	0	0	+	+	+
	15	0	0	0	0	0	-	0	-	-	-	0	0	0	0	0	0	0	-	-	0
	16	+	0	+	+	0	0	+	+	+	0	+	0	+	+	0	+	+	0	+	0
	17	0	0	0	0	0	0	0	0	+	+	0	0	0	0	0	0	0	0	+	+
	18	0	+	+	+	+	+	+	0	0	0	0	+	+	+	+	0	+	+	0	+

We validate the models on a single validation sample. In this way, the results are comparable. We report the type I, type II and unweighted error rates and the Gini-coefficient of the validation samples of the models 1 and 3 years prior to failure. Table 10 shows the results of the validation of the different models 1 year prior to failure and table 11 for the models 3 years prior to failure. The best models – based on the unweighted error rate – are highlighted.

Table 10: Results of the validation of the simple-intuitive models (SIM) 1 ypf

	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate	Gini-coefficient
Model 1	0.5763	28.46%	22.52%	25.49%	61.26%
Model 2	**0.5402**	**20.82%**	**27.98%**	**24.40%**	**63.50%**
Model 3	0.5795	25.04%	25.38%	25.21%	63.67%
Model 4	0.5786	27.48%	24.15%	25.82%	61.23%
Model 5	0.5484	26.51%	23.91%	25.21%	62.02%
Model 6	**0.5315**	**24.89%**	**23.31%**	**24.10%**	**62.31%**
Model 7	0.5149	34.79%	18.54%	26.66%	57.87%
Model 8	0.5177	32.08%	21.33%	26.70%	56.78%
Model 9	0.5234	29.33%	23.15%	26.24%	56.83%
Model 10	0.4923	35.73%	18.46%	27.10%	54.04%
Model 11	0.5726	27.79%	22.26%	25.02%	63.23%
Model 12	**0.5310**	**22.36%**	**26.15%**	**24.26%**	**64.24%**
Model 13	0.5758	24.32%	25.61%	24.97%	64.21%
Model 14	0.5755	26.66%	24.54%	25.60%	62.03%
Model 15	0.5488	24.17%	26.33%	25.25%	62.73%
Model 16	0.5712	27.71%	22.45%	25.08%	62.31%
Model 17	0.5556	25.94%	24.54%	25.24%	60.67%
Model 18	0.4989	28.43%	23.43%	25.93%	58.49%
Model 19	0.5599	28.20%	29.61%	28.91%	52.77%
Model 20	0.5004	32.27%	21.81%	27.04%	55.27%

Both 1 year prior to failure and 3 years prior to failure, the models 2, 6 and 12 are the best performing models. 3 years prior to failure, model 12 is clearly the best model, because it has the lowest UER and the highest Gini-coefficient. 1 year prior to failure however, model 6 has a lower UER. But, since this difference is very small and since model 12 has a much higher Gini-coefficient, we can conclude that, also 3 years prior to failure, model 12 is the best performing model.

This results in one model that performs best both on short and on medium term. This is a clear advantage over previous models (e.g. Ooghe-Joos-De Vos 1991) where two calculations had to be made to assess the financial health of a company. Here one model score suffices; it only has to be compared to two different cut-off points. The cut-off point 1 year prior to failure is always lower than the same point 3 years prior to failure.

Table 11: Results of the validation of the simple-intuitive models (SIM) 3ypf

	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate	Gini-coefficient
Model 1	0.6054	31.50%	42.14%	36.82%	36.29%
Model 2	**0.5532**	**32.38%**	**35.55%**	**33.97%**	**40.32%**
Model 3	0.6032	30.05%	39.93%	34.99%	39.95%
Model 4	0.6019	33.91%	38.32%	36.12%	37.35%
Model 5	0.568	35.32%	35.62%	35.47%	38.03%
Model 6	**0.5561**	**33.26%**	**35.16%**	**34.21%**	**37.70%**
Model 7	0.5806	34.97%	37.99%	36.48%	34.40%
Model 8	0.5621	35.82%	36.74%	36.28%	34.49%
Model 9	0.5618	33.59%	38.55%	36.07%	35.30%
Model 10	0.5804	37.57%	37.63%	37.60%	30.55%
Model 11	0.6029	29.46%	42.96%	36.21%	37.83%
Model 12	**0.5543**	**28.32%**	**38.63%**	**33.47%**	**41.49%**
Model 13	0.6026	27.01%	41.71%	34.36%	40.99%
Model 14	0.5996	31.88%	38.84%	35.36%	38.59%
Model 15	0.5661	33.11%	36.53%	34.82%	39.23%
Model 16	0.5944	35.04%	37.37%	36.20%	37.90%
Model 17	0.5857	32.14%	39.20%	35.67%	35.47%
Model 18	0.5416	30.02%	40.69%	35.36%	37.09%
Model 19	0.5802	29.19%	44.18%	36.69%	34.09%
Model 20	0.5442	33.41%	39.36%	36.38%	33.96%

5.3 Comparison to classical statistical methods

We also want to compare the performance results of the best simple-intuitive model (SIM 12) to the results of classical statistical models, both 1 year and 3 years prior to failure. As statistical models we use:

- the general linear model Ooghe-Verbaere 1982 (OV82);
- the conditional probability model Ooghe-Joos-De Vos 1991 (OJD 1ypf and OJD 3ypf);
- a new model with the variables of SIM 12, but now with coefficients based on linear regression (linear M 1ypf and linear M 3ypf);
- a new model with 8 ratios, produced by a forward stepwise logistic regression on the ratios 4 to 18 (logit M 1ypf and logit M 3ypf).

The composition and coefficients of the statistical models are shown in appendix 6. The performance results on the validation sample are compared in table 12 and table 13.

Table 12: Comparison of performance results of SIM 12 with the statistical models 1 ypf

	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate	Gini-coefficient
SIM 12	**0.5310**	**22.36%**	**26.15%**	**24.26%**	**64.24%**
OV82	0.1904	16.53%	29.44%	22.98%	66.33%
OJD91 1ypf	0.4142	23.98%	24.52%	24.25%	64.27%
Linear M 1ypf	0.9049	22.28%	24.87%	23.52%	66.06%
Logit M 1ypf	0.9332	25.49%	28.22%	26.85%	57.84%

Table 13: Comparison of performance results of SIM 12 with statistical models 3 ypf

	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate	Gini-coefficient
SIM 12	**0.5543**	**28.32%**	**38.63%**	**33.47%**	**41.49%**
OV82	0.3939	27.36%	35.07%	31.21%	46.77%
OVD91 3ypf	0.2797	24.74%	41.12%	32.93%	43.54%
Linear M 3ypf	0.7506	23.45%	39.75%	31.60%	46.36%
Logit M 3ypf	0.7269	34.25%	35.21%	34.73%	38.40%

In general the performance results may not seem too impressive compared to previous (international) models. However, this is due to a more realistic validation sample, that includes a large number of annual accounts of heterogeneous companies from all industries, sizes and ages[15].

The new simple-intuitive model 12 does not have systematically better or worse performance results in comparison to the more complex and less transparent statistical models. The "old" OV82 and the new linear models show better performance results although the difference is rather small.

Considering these results, we can state that our new simple-intuitive model is not secondary to the well-known statistical methods. On the contrary, this model combines comparable validation results with the advantages of more transparency and less complexity.

Conclusion

Most failure prediction models use statistical techniques such as multiple discriminant analysis and multiple logistic regression. Too often, the problems related to the use of statistical methods are neglected. In general, too complicated procedures reduce the stability and transparency and impose the problem of overfitting.

In this paper, a new type of failure prediction models was developed and tested, namely the simple-intuitive models. Eight ratios are first logit-transformed and then equally weighted to obtain a model score. The ratios are selected based on expertise, rather than on statistical techniques. Based on performance tests (the lowest

[15] See Ooghe and Balcaen (2002). In this paper is also shown that other international models do not necessarily give better results.

unweighted error rate and the highest Gini-coefficient) on a very extensive and rough data set, one model (SIM 12) scores best both on short term (1 year prior to failure) and on medium term (3 years prior to failure).
This new model was compared with different established and new statistical models. The performance results are comparable. Since the model does perform approximately equal, and it has the advantages of being simple, transparent and intuitively correct, we argue that it is superior to the well-known statistical models.
This paper provides a basis for future research on ‘simple-intuitive models’. For example, the different range of the variables and the treatment of special cases are methodological issues to be tackled. Also, the models can be expanded with additional, non-financial variables. A third idea is the construction of industry-specific models.

References

Altman E.I., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, Vol. 23, nr. 4, September 1968, p. 589-609

Altman E.I., Narayanan P., 1997, An international survey of business failure classification models. Financial Markets, Institutions and Instruments, Vol. 6, nr. 2, p. 1-57

Balcaen S., Ooghe H., 2004, 35 years of studies on business failure: an overview of the classical statistical methodologies and their related problems, Paper nr. 04/248, Working paper series, Faculty of Economics and Business Administration, Ghent University, Belgium, 56 pp. (accepted for publication in British Accounting Review after substantial reduction)

Beaver W., 1967, Financial ratios predictors of failure. Empirical Research in Accounting: Selected Studies 1966, Journal of Accounting Research, Supplement to Vol. 4, p. 71-111

Bell T.B., Ribar G.S., Verchio J.R., 1990, Neural nets vs. logistic regression: a comparison of each model's ability to predict commercial bank failures, paper submitted to 'Cash Flow Accounting Conference' (Nice)

Bilderbeek J., 1979, An empirical study of the predictive ability of financial ratios in the Netherlands. Zeitschrift Für Betriebswirtschaft, May 1979, p. 388-407

Curram S.P., Mingers J., 1994, Neural networks, decision tree induction and discriminant analysis: an empirical comparison, Journal of the Operational Research Society, Vol. 45, nr. 4, April 1994, p. 440-450

Doumpos M., Zopoudinis C., 1999, A multicriteria discrimination method for the prediction of financial distress: the case of Greece. Multinational Finance Journal, Vol. 3, nr. 2, p. 71-101

Gloubos G., Grammatikos T., 1988, The success of bankruptcy prediction models in Greece. Studies in Banking and Finance, Vol. 7, p. 37-46

Hand D. J., Marginal classifier improvement and reality, presented on 'Symposium on data mining' (Ghent, May 10, 2004), <http://allserv.ugent.be/~wdewolf/CVStat/Symposia/DataMining/SlidesDHand.pdf>

Joos Ph., Ooghe H., Sierens N., 1998, Methodologie bij het opstellen en beoordelen van kredietclassificatiemodellen. Tijdschrift voor Economie en Management, Vol. 18, nr. 1, p. 1-48

Karels G.V., Prakash A.J., 1987, Multivariate normality and forecasting of business bankruptcy. Journal of Business Finance & Accounting, Vol. 14, nr. 4, Winter 1987, p. 573-593

Keasey K., McGuinness P., Short H., 1990, Multilogit approach to predicting corporate failure: Further analysis and the issue of signal consistency. Omega International Journal of Management Science, Vol. 18, nr. 1, p. 85-94

Koh H.C., 1992, The sensitivity of optimal cutoff points to misclassification costs of Type I and Type II errors in the going-concern prediction context. Journal of Business Finance & Accounting, Vol. 19, nr. 2, January 1992, p. 187-197

Laitinen T., Kankaanpää M., 1999, Comparative analysis of failure prediction methods: the Finnish case. The European Accounting Review, Vol. 8, nr. 1, p. 67-92

Martin D., 1977, Early warning of bank failure: a logit regression approach, Journal of Banking & Finance, Vol. 1, nr. 2/3, p. 249-276

Mossman Ch.E., Bell G.G., Swartz L.M., Turtle H., 1998, An empirical comparison of bankruptcy models. The Financial Review, Vol. 33, nr. 2, p. 35-54

Ohlson J., 1980, Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, Vol. 18, nr. 1, Spring 1980, p. 109-131

Ooghe H., Balcaen S., 2002, Are failure prediction models transferable from one country to another? An empirical study using Belgian financial statements, Paper nr. 02/132, Working Paper Series, Faculty of Economics and Business Administration, Ghent University, Belgium, 42 pp

Ooghe H., Joos P., De Bourdeaudhuij C., 1995, Financial distress models in Belgium: The results of a decade of empirical research. International Journal of Accounting, Vol. 30, p. 245-274

Platt H.D., Platt M.B., 1990, Development of a class of stable predictive variables: the case of bankruptcy prediction. Journal of Business Finance & Accounting, Vol. 17, nr. 1, Spring 1990, p. 31-51

Steele A., 1995, Going concern qualifications and bankruptcy prediction. Paper presented at the 18th Annual Congress of the European Accounting Association, 10-12 May, Birmingham, UK, p. 1-28

Zavgren C.V., 1985, Assessing the vulnerability to failure of American industrial firms: A logistic analysis. Journal of Business Finance and Accounting, Vol. 12, nr. 1, Spring 1985, p. 19-45

Appendix 1: Calculation of the ratios

In Belgium, companies are required to deposit their annual accounts in a prescribed form, dependent on their size. A distinction is made between ‘large’ companies that must prepare their annual accounts in a **complete form**, and ‘small’ companies that are allowed to prepare their annual accounts in an **abbreviated form**.

The group of larger companies consists of companies with more than 100 employees, plus companies that meet at least two of three criteria concerning number of employees (≥ 50 employees), turnover (≥ 625 000 euro) and total assets (≥ 3 125 000 euro). A major percentage of the companies have annual accounts in an abbreviated form.

The complete form annual accounts have a slightly different, but more extensive format than the abbreviated form annual accounts. Each of these forms uses different **codes**. The codes mentioned in this table, refer to the codes that are reported in the financial statements.

“<>” means that the amount mentioned under a certain section can either be positive or negative and that the sign has to be taken into account. For all the other codes the numbers have to be considered in absolute value (without positive or negative sign) and added or subtracted according to the + or – sign in the formula.

Ratio		**Annual account sections Complete form**	**Annual account sections Abbreviated form**
1	Gross added value / personnel employed	(70/74 – 740 – 60 – 61) / 9087	(70/61 – 61/70) / 9087
2	Personnel charges / personnel employed	<62> / 9087	<62> / 9087
3	Gross added value / value of production	(70/74 – 740 – 60 – 61) / (70/74 – 740)	(70/61 – 61/70) / (70/61 – 61/70 + 60/61)
4	Gross added value / personnel charges	(70/74 – 740 – 60 – 61) / <62>	(70/61 – 61/70) / <62>
5	Financial leverage	[(70/66 - 66/70 + 780 – 680 - <65> – 9126) / 20/58] - [(- <65> - 9126 – 6560 + 6561) / (17 + 42/48)]	[(70/66 - 66/70 + 780 – 680 - <65> – 9126) / 20/58] - [(- <65> - 9126 – <656>) / (17 + 42/48)]
6	Net return on total assets before taxes	(70/67 - 67/70 + 650 + 653 – 9126 + 9134) / 20/58	(70/66 – 66/70 + 780 – 680 - <65> - 9126 - <656>) / 20/58
7	Net return on equity after taxes	(70/67 – 67/70) / <10/15>	(70/67 - 67/70) / <10/15>
8	Gross return on equity after taxes	(70/67 - 67/70 + 630 + <631/4> + <635/7> + 6501 + <651> + 6560 – 6561 + 660 + 661 + <662> + 663 + 680 - 760 – 761 – 762 – 780 – 9125) / <10/15>	(70/67 – 67/70 + <656> - 780 + 680 + 8079 – 8089 + 8279 – 8289 + 8475 – 8485 - <631/4> - <635/7> - 9125) / <10/15>
9	Self-financing level	(13 + 140 – 141) / 10/49	(13 + 140 – 141) / 10/49
10	General level of financial independence	<10/15> / 10/49	<10/15> / 10/49
11	Debts guaranteed / total debt	(9061 + 9062) / (17 + 42/48)	(9061 + 9062) / (17 + 42/48)
12	Short term financial debt level	430/8 / 42/48	430/8 / 42/48

13	Cash flow after taxes / liabilities	(70/67 - 67/70 + 630 + <631/4> + <635/7> + 6501 + <651> + 6560 – 6561 + 660 + 661 + <662> + 663 + 680 - 760 – 761 – 762 – 780 – 9125) / (16 + 17/49)	(70/67 – 67/70 + <656> - 780 + 680 + 8079 – 8089 + 8279 – 8289 + 8475 – 8485 - <631/4> - <635/7> - 9125) / (16 + 17/49)
14	Free cash flow / financial debt	(70/67 - 67/70 + 630 + <631/4> + <635/7> + 6501 + <651> + 6560 – 6561 + 660 + 661 + <662> + 663 + 680 - 760 – 761 – 762 – 780 – 9125 – 802 + 803 + 809 – 810 –816 + 817 – 822 + 823 + 829 – 830 – 836 + 837 – 842 + 843 + 849 – 850 + <854> - 858 + 859) / (170/4 + 42 + 43)	(70/67 – 67/70 + <656> - 780 + 680 + 8079 – 8089 + 8279 – 8289 + 8475 – 8485 - <631/4> - <635/7> - 9125 – 8029 – 8169 – 8365 + 8039 + 8179 + 8375 - 8229 - 8425 + 8239 + 8435 + 8099 + 8299 + 8495 – 8109 – 8309 – 8505 + <8545>) / (170/4 + 42 + 43)
15	Overdue taxes and social security charges / taxes, remuneration and social security debt	(9072 + 9076) / 45	(9072 + 9076) / 45
16	Current ratio	(29/58 – 29) / (42/48 – 492/3)	(29/58 – 29) / (42/48 – 492/3)
17	(Cash + short-term investments) / total assets	(51/53 + 54/58) / 20/58	(50/53 + 54/58 – 8721) / 20/58
18	(Cash + short-term investments - financial debt) / current assets	(50/53 + 54/58 – 43) / (29/58 – 29)	(50/53 – 54/58 – 43) / (29/58 – 29)

Appendix 2: Overview of total numbers of companies and annual accounts (in population and samples)

	Total	***Non-failing***			***Failing***		
		Total	Pure	Doubt-causing	Total	Pure	Doubt-causing
Population of companies 1990-2001	322,976	292,003	252,496	39,507	30,973	30,664	309
Sample of companies 1990-2001							
- Estimation	57,574	42,226	42,226	0	15,348	15,348	0
- Validation	65,491	49,948	41,945	8,003	15,543	15,316	227
Sample of annual accounts 1990-1999							
- Estimation 1 ypf	51,101	39,609	39,609	0	11,492	11,492	0
- Estimation 3 ypf	53,760	39,609	39,609	0	14,151	14,151	0
- Validation 1 ypf	58,838	47,310	39,377	7,933	11,528	11,440	88
- Validation 3 ypf	61,673	47,310	39,377	7,933	14,363	14,235	128
Available annual accounts 1990-1999							
- Estimation 1 ypf	30,849	27,898	27,898	0	2,591	2,591	0
- Estimation 3 ypf	38,420	27,898	27,898	0	10,522	10,522	0
- Validation 1 ypf	34,622	31,946	27,799	4,147	2,676	2,615	61
- Validation 3 ypf	42,570	31,946	27,799	4,147	10,624	10,536	88
Available annual accounts 1990-1999 excluding ratio 5							
- Estimation 1 ypf	30,133	27,565	27,565	0	2,568	2,568	0
- Estimation 3 ypf	37,997	27,565	27,565	0	10,432	10,432	0
- Validation 1 ypf	34,078	31,422	27,476	3,946	2,656	2,595	61
- Validation 3 ypf	41,932	31,422	27,476	3,946	10,510	10,422	88

Appendix 3: List of ratios arranged by discriminating power

Listed below are the logit values of the ratios arranged by their discriminating power, i.e. their D-max. D-max can be defined as the largest difference between the cumulative distribution function of the scores of the failing firms ($F_{failing}(y)$) and the cumulative distribution function of the scores of the non-failing firms ($F_{non\text{-}failing}(y)$). We list the largest positive differences (D_{pos}-max) and the largest negative differences (D_{neg}-max). D-max is the highest absolute value.
Important to mention is that the ratios 2, 12 and 15 have an (expected) opposite sign, so the D-max is originally negative. Therefore, we expect the D-max of these ratios to be the absolute value of D_{neg}-max. However, this is not true for ratio 2. This can be explained by the fact that better-performing companies (with higher added value) often pay more and thus have higher personnel costs than failing companies.

Estimation sample: non-failing versus failing - 1 year prior to failure

	D_{pos}-max	D_{neg}-max	D-max	Gini
L_7	53.4543	-1.9333	53.4333	60.72%
L_{10}	49.5234	0	49.5234	63.05%
L_8	49.3468	-4.4723	49.3468	46.99%
L_{13}	48.6640	-0.2607	48.6640	61.49%
L_9	48.4836	-0.0063	48.4836	60.50%
L_5	44.9076	-0.4110	44.9076	54.88%
L_6	43.7047	-0.4697	43.7047	52.36%
L_4	42.4409	0	42.4409	50.58%
L_{18}	37.7159	0	37.7159	44.17%
L_{16}	34.7711	0	34.7711	39.92%
L_{14}	31.4535	0	31.4535	34.88%
L_{12}	0.1395	-30.4644	30.4644	-13.70%
L_{15}	0	-26.0653	26.0653	-25.59%
L_{17}	24.7690	-0.2356	24.7690	31.50%
L_{11}	3.4635	-5.5660	5.5660	49.44%

Estimation sample: non-failing versus failing - 3 years prior to failure

	D_{pos}-max	D_{neg}-max	D-max	Gini
L_{13}	37.0767	-0.2615	37.0767	43.32%
L_{10}	31.1543	0	31.1543	41.41%
L_{18}	29.8036	-0.3453	29.8036	36.42%
L_{9}	29.5173	0	29.5173	39.21%
L_{7}	28.8348	-2.1070	28.8348	31.29%
L_{8}	24.8193	-4.2934	24.8193	18.90%
L_{12}	0.1751	-24.5805	24.5805	-5.06%
L_{4}	24.4752	0	24.4752	30.74%
L_{6}	21.5502	-0.2150	21.5502	26.58%
L_{16}	21.5081	-0.6655	21.5081	22.51%
L_{5}	21.4271	-0.2013	21.4271	29.28%
L_{14}	20.6767	0	20.6767	23.86%
L_{17}	18.7596	0	18.7596	25.12%
L_{15}	0	-14.2092	14.2092	-14.00%
L_{11}	2.6929	-1.8750	2.6929	55.29%

Appendix 4: Correlation matrices of the ratios

Estimation sample 1 year prior to failure: Spearman correlation (≥ 0.60 in bold)

		RATIO4	RATIO5	RATIO6	RATIO7	RATIO8	RATIO9	RATIO10	RATIO11	RATIO12	RATIO13	RATIO14	RATIO15	RATIO16	RATIO17	RATIO18
RATIO4	Correlation coefficient	**1.000**														
	Sig. (2tailed)	.														
RATIO5	Correlation coefficient	0.383	**1.000**													
	Sig. (2tailed)	0.000	.													
RATIO6	Correlation coefficient	0.435	**0.890**	**1.000**												
	Sig. (2tailed)	0.000	0.000	.												
RATIO7	Correlation coefficient	0.373	**0.874**	**0.855**	**1.000**											
	Sig. (2tailed)	0.000	0.000	0.000	.											
RATIO8	Correlation coefficient	0.380	0.594	**0.602**	**0.602**	**1.000**										
	Sig. (2tailed)	0.000	0.000	0.000	0.000	.										
RATIO9	Correlation coefficient	0.192	0.491	0.498	0.453	0.137	**1.000**									
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	.									
RATIO10	Correlation coefficient	0.174	0.268	0.287	0.287	-0.123	**0.667**	**1.000**								
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	.								
RATIO11	Correlation coefficient	0.055	-0.029	0.021	-0.010	0.081	-0.021	-0.124	**1.000**							
	Sig. (2tailed)	0.000	0.000	0.000	0.076	0.000	0.000	0.000	.							
RATIO12	Correlation coefficient	0.091	0.150	0.066	0.123	0.021	0.152	0.250	-0.147	**1.000**						
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.						
RATIO13	Correlation coefficient	0.424	**0.692**	**0.719**	**0.656**	0.568	0.561	0.524	-0.046	0.197	**1.000**					
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.					
RATIO14	Correlation coefficient	0.293	0.501	0.524	0.484	0.349	0.344	0.295	-0.051	0.102	0.575	**1.000**				
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.				
RATIO15	Correlation coefficient	0.078	0.071	0.060	0.078	0.030	0.094	0.115	-0.014	0.085	0.072	0.045	**1.000**			
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000	.			
RATIO16	Correlation coefficient	0.033	0.238	0.252	0.237	-0.067	0.528	**0.628**	-0.108	0.186	0.362	0.277	0.064	**1.000**		
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.		
RATIO17	Correlation coefficient	0.002	0.240	0.202	0.208	0.064	0.288	0.290	-0.160	0.420	0.296	0.192	0.046	0.401	**1.000**	
	Sig. (2tailed)	0.770	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	
RATIO18	Correlation coefficient	0.151	0.204	0.158	0.184	0.050	0.256	0.361	-0.127	0.704	0.303	0.152	0.087	0.278	**0.710**	**1.000**
	Sig. (2tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.

Estimation sample 3 years prior to failure: Spearman correlation (≥ 0.60 in bold)

		RATIO4	RATIO5	RATIO6	RATIO7	RATIO8	RATIO9	RATIO10	RATIO11	RATIO12	RATIO13	RATIO14	RATIO15	RATIO16	RATIO17	RATIO18
RATIO4	Correlation coefficient	**1,000**														
	Sig. (2-tailed)	,														
RATIO5	Correlation coefficient	0,420	**1,000**													
	Sig. (2-tailed)	0,000	,													
RATIO6	Correlation coefficient	0,473	**0,890**	**1,000**												
	Sig. (2-tailed)	0,000	0,000	,												
RATIO7	Correlation coefficient	0,402	**0,872**	**0,848**	**1,000**											
	Sig. (2-tailed)	0,000	0,000	0,000	,											
RATIO8	Correlation coefficient	0,410	**0,609**	**0,615**	**0,610**	**1,000**										
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	,										
RATIO9	Correlation coefficient	0,228	0,492	0,499	0,468	0,165	**1,000**									
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	,									
RATIO10	Correlation coefficient	0,189	0,281	0,301	0,326	-0,074	**0,677**	**1,000**								
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	,								
RATIO11	Correlation coefficient	0,059	-0,022	0,028	-0,012	0,074	-0,004	-0,101	**1,000**							
	Sig. (2-tailed)	0,000	0,000	0,000	0,019	0,000	0,409	0,000	,							
RATIO12	Correlation coefficient	0,071	0,143	0,052	0,120	0,020	0,149	0,241	-0,146	**1,000**						
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,						
RATIO13	Correlation coefficient	0,477	**0,700**	**0,721**	**0,666**	**0,600**	0,549	0,504	-0,031	0,198	1,000					
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,					
RATIO14	Correlation coefficient	0,327	0,514	0,542	0,496	0,385	0,335	0,284	-0,026	0,080	0,584	**1,000**				
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,				
RATIO15	Correlation coefficient	0,057	0,051	0,038	0,054	0,007	0,081	0,105	-0,020	0,083	0,067	0,033	**1,000**			
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,152	0,000	0,000	0,000	0,000	0,000	0,000	,			
RATIO16	Correlation coefficient	0,053	0,241	0,256	0,255	-0,046	0,524	**0,624**	-0,094	0,175	0,332	0,270	0,056	**1,000**		
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,		
RATIO17	Correlation coefficient	0,012	0,210	0,174	0,190	0,056	0,260	0,276	-0,148	0,422	0,276	0,164	0,056	0,361	**1,000**	
	Sig. (2-tailed)	0,015	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,	
RATIO18	Correlation coefficient	0,132	0,198	0,145	0,185	0,052	0,259	0,358	-0,127	**0,737**	0,303	0,137	0,091	0,279	**0,710**	1,000
	Sig. (2-tailed)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	,

Appendix 5: Results of the models with the *x* most discriminating ratios

Below are the error rates for different models containing a different numbers of ratios. The model score is calculated by equation (4):

$$S_j = \frac{\sum_{i=1}^{n} L_{ij}}{n}$$

The order of the ratios is given by their discriminating power, which can be found in Appendix 3. The results show that models of more than 8 ratios are not expected to give significantly better results than models with 8 ratios or less.
Surprising is that the optimal number of ratios is relatively low: the best results are made with only 2 to 3 ratios 1 year prior to failure and 3 to 4 ratios 3 years prior to failure. This can be explained by the fact that adding (less discriminating) ratios diminishes the relative weight of the most discriminating ratios.

1 year prior to failure

Number of ratios	Ratio added	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate
1	L_7	0.4503	27.15%	22.66%	24.90%
2	L_{10}	0.5009	26.51%	21.70%	24.10%
3	L_8	0.5181	25.72%	22.50%	24.11%
4	L_{13}	0.4952	30.95%	17.89%	24.42%
5	L_9	0.5126	24.13%	24.82%	24.48%
6	L_5	0.4838	32.00%	17.04%	24.52%
7	L_6	0.4921	30.20%	19.07%	24.63%
8	L_4	0.5456	25.04%	24.80%	24.92%
9	L_{18}	0.5363	25.04%	24.08%	24.56%
10	L_{16}	0.5581	24.40%	25.39%	24.89%
11	L_{14}	0.5403	27.15%	24.79%	25.97%
12	L_{12}	0.5277	28.77%	22.65%	25.71%

3 years prior to failure

Number of ratios	Ratio added	Cut-off point	Type I error rate	Type II error rate	Unweighted error rate
1	L_{13}	0.5203	31.39%	40.77%	36.08%
2	L_{10}	0.5375	25.06%	39.96%	32.51%
3	L_{18}	0.5287	23.87%	37.94%	30.91%
4	L_9	0.5275	19.43%	41.99%	30.71%
5	L_7	0.5202	28.34%	37.43%	32.89%
6	L_8	0.5281	34.30%	34.63%	34.46%
7	L_{12}	0.5243	31.28%	37.20%	34.24%
8	L_4	0.5625	32.16%	36.47%	34.31%
9	L_6	0.5552	33.56%	35.48%	34.52%
10	L_{16}	0.5785	31.15%	38.45%	34.80%
11	L_5	0.5672	34.37%	35.38%	34.88%
12	L_{14}	0.5583	34.02%	36.58%	35.02%

Appendix 6: Composition of models OV82, OJD 91, linear M and logit M

OV82: Ooghe-Verbaere 1982

The tables below illustrate the composition of the general OV model one to three years prior to failure, reporting the included variables and the non-standardized coefficients of the linear discriminant model.

OV 82 general model 1 to 3 years prior to failure

	Variables	**Codes complete form**	**Codes abbreviated form**	**Non-standardized coefficients**
	Intercept	-	id.	+ 0.2324
X1	(Retained earnings + accumulated profits or losses) / (Equity + liabilities)	(|13| + |140| - |141|) / (|10/49|)	id.	+ 4.3178
X2	Overdue taxes and social security debt / Short-term debt	(|9072| + |9076|) / (|42/48| + |492/3|)	id.	- 11.6782
X3	Liquid assets / Restricted current assets	(|54/58|) / (|29/58| - |29|)	id.	+ 3.1676
X4	(Work in progress, finished goods and contracts in progress) / Current working assets	(|32| + |33| + |37|) / (|3| + |40/41| + |490/1|)	(|3|) / (|3| + |40/41| + |490/1|)	- 1.6200
X5	Short-term financial debt to credit institutions / Short-term debt	(|430/8|) / (|42/48| + |429/3|)	id.	- 0.8353

OJD91: Ooghe-Joos-Devos 1991

The tables below illustrate the composition of the OJD models one and three years prior to failure, reporting the included variables of the logistic regression. Also the codes are listed. The coefficients cannot be indicated, because of an exclusive licence contract with Graydon NV.

OJD 91 1 year prior to failure

<table>
<tr><th></th><th>Variables</th><th>Codes complete form</th><th>Codes abbreviated form</th></tr>
<tr><td>X1</td><td>Direction of the financial leverage = Net return on total assets before taxes – average interest rate of debt (1 if > 0, 0 if <0)</td><td>{(|70/66| - |66/70| + |780| - |680| - <65> -|9126|) / |20/58|} - {(-<65> - |9126| - |6560| + |6561|) / (|17| + |42/48|)}</td><td>{(|70/66| - |66/70| + |780| - |680| - <65> -|9126|) / |20/58|} - {(-<65> - |9126| - <656>) /(|17| + |42/48|)}</td></tr>
<tr><td>X2</td><td>(Accumulated profits or losses + retained earnings) / Equity and total liabilities less accrued charges and deferred income</td><td>(|13| + |140| - |141|) / (|10/49| - |492/3|)</td><td>id.</td></tr>
<tr><td>X3</td><td>Cash and short-term investments / Total assets</td><td>(|51/53| + |54/58|) / |20/58|</td><td>(|50/53| + |54/58| - |8721|) / |20/58|</td></tr>
<tr><td>X4</td><td>Overdue taxes and social security debt (1 if >0, 0 else)</td><td>(|9072| + |9076|) 1 if >0, else 0</td><td>id.</td></tr>
<tr><td>X5</td><td>(Inventories + accounts receivable – accounts payable – taxes, remuneration and social security debt – advances received on contracts in progress) / Total assets</td><td>(|3| + |40/41| - |44| - |45| - |46|) / (|20/58|)</td><td>id.</td></tr>
<tr><td>X6</td><td>Net return on operating assets before taxes</td><td>(|70/64| - |64/70| + |9125|) / (|20| +|21| + |22/7| + |3| + |40/41|)</td><td>id.</td></tr>
<tr><td>X7</td><td>Short-term financial debt / Short-term debt</td><td>(|430/8|) / (|42/48|)</td><td>id.</td></tr>
<tr><td>X8</td><td>Debts guaranteed / Total debt</td><td>(|9061| + |9062|) / (|17| + |42/48|)</td><td>id.</td></tr>
</table>

OJD 91 3 years prior to failure

<table>
<tr><th></th><th>Variables</th><th>Codes complete form</th><th>Codes abbreviated form</th></tr>
<tr><td>X1</td><td>(Accumulated profits or losses + retained earnings) / Equity and total liabilities less accrued charges and deferred income</td><td>(|13| + |140| - |141|) / (|10/49| - |492/3|)</td><td>id.</td></tr>
<tr><td>X2</td><td>Publication lag of the annual accounts (in days)</td><td>-</td><td>-</td></tr>
<tr><td>X3</td><td>Overdue taxes and social security debt (1 if >0, 0 else)</td><td>(|9072| + |9076|) 1 if >0, else 0</td><td>id.</td></tr>
<tr><td>X4</td><td>(Earnings before interest, taxes, depreciation and amortization (EBITDA) – capital investments) /Total assets</td><td>{(|70/66|- |66/70| - <65> - |9126| - <631/4> + <635/7> + |807| - |808| + |827 | - |828| + |847| -|848| - |860| - |861| - |9125|) - (|816| - |817| + |822| - |823| - |829| + |830| + |836| - |837| + |842| +| 843| -|849| + |850| - <854> + |858| - |859|)} / (|20/58|)</td><td>{(|70/66|- |66/70| - <65> - |9126| - <631/4> - <635/7> + |8079| - |8089| + |8279| - |8289| + |8475| - |8485| - |9125|) - (|8169| - |8179| + |8229| - |8239| - |8299| + |8309| + |8365| - |8375| + |8425| - |8435| - |8495| + |8505| - <8545>)} / (20/58)</td></tr>
<tr><td>X5</td><td>Relationships with affiliated enterprises = (amounts receivable from them + commitments guaranteed on their behalf + other financial commitments in their favour) / Total assets</td><td>(|9291| + |9381| + |9401|) / (|20/58|)</td><td>(|9291| + |9294| + |9295|) / (|20/58|)</td></tr>
<tr><td>X6</td><td>Total debt / Equity and total liabilities less accrued charges and deferred income</td><td>(|17| + |42/48|) / (|10/49| - |492/3|)</td><td>id.</td></tr>
</table>

Linear M: best simple-intuitive model 12 with coefficients based on linear regression

The ratios of model 12 – the best performing intuitive model – are used as independent variables in a linear regression (or discriminant analysis) on the estimation samples 1ypf and 3ypf after logit transformation. The dependent variables are 1 (non-failing) and 0 (failing). All coefficients are significant at the 0.05 level, except ratio 13 1 ypf.

1 year prior to failure

	Unstandarized coefficient	**Sig.**
(Constant)	0.24293	0.000
L_4	0.05581	0.000
L_6	0.14908	0.000
L_7	0.22098	0.000
L_9	0.06548	0.002
L_{10}	0.46370	0.000
L_{12}	0.25977	0.000
L_{13}	-0.01396	0.362
L_{18}	0.08530	0.000

3 years prior to failure

	Unstandarized coefficient	**Sig.**
(Constant)	-0.27121	0.000
L_4	0.10341	0.000
L_6	-0.42177	0.000
L_7	0.21952	0.000
L_9	0.10983	0.000
L_{10}	0.61995	0.000
L_{12}	0.68179	0.000
L_{13}	0.32698	0.000
L_{18}	0.26677	0.000

Logit M: model with 8 ratios based on forward stepwise logit-analysis

A logit analysis is applied on the estimation samples 1ypf and 3 ypf before logit transformation. The independent variables are 1 (non-failing) and 0 (failing); the independent variables are the ratios 4 to 18. Forward stepwise is used until 8 ratios are selected. All coefficients are significant at the 0.05 level.

1 year prior to failure

	Coefficient	**Sig.**
(Constant)	2.63770	0.000
Ratio 4	0.00044	0.000
Ratio 7	0.00159	0.000
Ratio 8	0.00053	0.000
Ratio 9	-0.00053	0.027
Ratio 10	0.00318	0.003
Ratio 11	0.29180	0.003
Ratio 12	-1.07195	0.000
Ratio 17	2.55522	0.000

3 years prior to failure

	Coefficient	**Sig.**
(Constant)	0.92397	0.000
Ratio 4	0.00049	0.000
Ratio 7	0.00073	0.000
Ratio 8	0.00016	0.000
Ratio 11	0.43073	0.000
Ratio 12	-1.28600	0.000
Ratio 14	0.00008	0.001
Ratio 17	1.94088	0.000
Ratio 18	-0.00083	0.000