Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time

du Jardin, Philippe and Severin, Eric

EDHEC Business School

28 June 2011

Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time

Philippe du Jardin[a,1,*], Eric Séverin[b,2]

[a]*Edhec Bussiness School, 393 Promenade des Anglais, BP 3116, 06202 Nice Cedex 3, France*
[b]*Université de Lille 1, 104 Avenue du Peuple Belge, 59000 Lille, France*

Abstract

This study attempts to show how a Kohonen map can be used to improve the temporal stability of the accuracy of a financial failure model. Most models lose a significant part of their ability to generalize when data used for estimation and prediction purposes are collected over different time periods. As their lifespan is fairly short, it becomes a real problem if a model is still in use when re-estimation appears to be necessary. To overcome this drawback, we introduce a new way of using a Kohonen map as a prediction model. The results of our experiments show that the generalization error achieved with a map remains more stable over time than that achieved with conventional methods used to design failure models (discriminant analysis, logistic regression, Cox's method, and neural networks). They also show that type-I error, the economically costliest error, is the greatest beneficiary of this gain in stability.

Keywords: decision support systems, finance, bankruptcy prediction, self-organizing map

1. Introduction

Models that have long been used by banks and rating agencies to forecast firm failure, have many drawbacks that have given rise to an extensive body of literature (Balcaen and Ooghe, 2006). Nearly all of the drawbacks (whether related to modelling techniques, sampling and variable selection procedures, control parameters, model design or validation processes) that could have an effect on their robustness have been analysed. But one of these drawbacks, having to do with data stationarity, has not been overcome. A forecasting model relies on the assumption that the relationship between the dependent variable (i.e. failure probability) and all independent variables

---

*Corresponding author
Email addresses:* `philippe.dujardin@edhec.edu` (Philippe du Jardin), `e.severin@wanadoo.fr` (Eric Séverin)
[1]Tel.: +33 (0)4 93 18 99 66 - Fax: +33 (0)4 93 83 08 10
[2]Tel.: +33 (0)3 20 12 29 99

is stable over time (Zavgren, 1983). Yet there is evidence that this stability is highly questionable (Charitou et al., 2004) and that the true forecasts of a model may be weak if this assumption is not fulfilled (Mensah, 1984). Indeed, models are sensitive to some parameters that describe macro-economic environments, and any change may influence their accuracy (Mensah, 1984; Platt et al., 1994). In practice, then, models need to be re-estimated frequently to counterbalance the effects of such phenomena (Grice and Ingram, 2001). However, nobody knows what their life span is, how often they need to be re-estimated. This uncertainty has a cost, the cost of the error made when a model unexpectedly enters its instability zone, and especially the cost of type-I errors, that is, the cost of predicting that a firm will survive when in fact it will go bankrupt. In such circumstances, the potential cost for an investor or a creditor who decides, for example, to lend money based on a bankruptcy risk probability involves a net loss in capital that will not be reimbursed, whereas type-II errors involve only the loss of a commercial bargain. For these reasons, we study a means to improve model stability over time.

Two main parameters lie at the root of model instability when there is a change in the economic environment between the period during which a model is estimated and that during which it is used for prediction. Firstly, the boundary that makes it possible to discriminate between healthy and unsound companies moves slightly (Pompe and Bilderbeek, 2005). Secondly, the distribution of explanatory variables changes (Pinches et al., 1973); variable mean and standard deviation are no longer identical and this phenomenon influences model accuracy. In this study, we focus on the latter issue so as to mitigate the effect of sampling variations. Instead of using financial indicators as explanatory variables, we proposed using them in a different way: these financial variables were used to design a set of regions at risk and to compute the ways companies moved within these regions over time. These moves were then quantified to represent standard behaviours, called "trajectories", and these trajectories were used to make forecasts. We thus developed a typology of behaviours, some leading to bankruptcy, others not, and we studied both their forecasting ability and their ability to provide estimates less sensitive to macro-economic changes than those of traditional models.

Regions at risk and trajectories were designed using Kohonen maps, and the prediction ability of trajectories was compared to that of models designed using discriminant analysis, logistic regression, Cox's method and a neural network[3]. We made these comparisons with data collected over different time periods, experiencing various economic conditions; models were estimated with data collected

---

[3]The neural network used in this study is a multilayer Perceptron.

over periods of either economic growth or downturn, and their prediction ability was assessed with data collected over similar or dissimilar periods.

The remainder of this paper is organised as follows. In section 2, we present a literature review that explains our research question. In section 3, we describe the samples and methods used in our experiments. In section 4, we present and discuss our results and in section 5, we conclude and suggest further research.

2. Literature review

Most financial failure prediction models rely on regression or classification techniques and were designed with single-period data. A model makes it possible to forecast the fate of a company at time t depending on data measured at time $t-1$. It is therefore assumed that, between the point in time when the regression or classification function is estimated and that when the function is to be used for a prediction purpose, the relationship between a probability of failure and variables used to compute it (financial ratios, for the most part) is stable (Mensah, 1984). It is also assumed that the extent to which variables are correlated does not change (Zavgren, 1983). However, it has been shown that these assumptions do not hold (Altman and Eisenbeis, 1978). Indeed, both the relationship between the dependent and independent variables and the distribution of explanatory variables are likely to be influenced by macro-economic phenomena. Therefore, any change in environmental conditions may greatly reduce model accuracy. It has been demonstrated that variations in economic cycles (alternating periods of economic growth and downturn or recession) and, to a lesser extent, changes that firms may face in terms of interest rates, credit policy, tax rates, competitive structures, technological cycles and institutional environment, have an influence on financial ratio distributions and on the boundary between failed and non-failed companies. This influence may result in models having poor prediction abilities (Mensah, 1984; Platt et al., 1994; Grice and Dugan, 2003). Of course, other parameters may play a role, especially when models are used with data that are outside their scope of validity. This is the case when a model is designed for a given firm's size (or for a particular sector or country) and is used with companies that do not meet these criteria. But the latter parameters are easily monitored and controlled, unlike the former.

Although the factors discussed here do not exhaust all possible explanations, the consequences of what has just been described may partly explain the data presented in Tables 1 and 2. Table 1 presents the main financial failure models whose generalization ability was studied in the financial

3

literature. This table includes Altman's (1968), Wilcox's (1973), Ohlson's (1980), Taffler's (1983) and Zmijewski's (1984) models, which are among the most popular. It also shows the correct classification rates of healthy and failed companies that are achieved with each model, including the sample size used to test model accuracy, companies' sectors of operations, and the time period during which data were collected.

Table 2 shows the different studies that assessed the prediction performance of all these models under the same conditions as those used when they were designed (identical modelling technique, identical sector), except the time period during which data were collected. And models were evaluated either in their original form or in a re-estimated form. All but two studies achieved correct classification rates far lower than those reported by their initial authors. Altman's (1968) model, which originally had an accuracy rate of 95.5%, ultimately had an accuracy rate ranging from 85% to 89.4% with five studies, and less than 80% with thirteen others. The results achieved with other models are similar. However, when original model coefficients are re-estimated to cope with the characteristics of the period during which they were used once again, models achieved rather better results. Four of the five studies that use Altman's (1968) model, and that achieve the best results, managed to do so with a re-estimated function. The same conclusion can be drawn from the results achieved with Ohlson's (1980) model. As far as the others are concerned, except Barth et al.'s (1985) model, used by Elmer and Borowski (1988), and Hopwood et al.'s (1994) model, used by McKee and Greenstein (2000), they all lead to the conclusion that the difference between the performance of a model, as originally reported, and the performance achieved, with a new sample, is fairly clear. These tables show clearly why failure models must be re-estimated frequently. They also show the limits of traditional validation procedures, where models are tested with samples collected within the same timeframe as that used to estimate them. However, no one knows how often such a re-estimation is to be done, even if it is clear that the more complex a classifier, the more often it must be re-estimated (Finlay, 2011). This fact is not without consequences for banking and financial institutions when some of their decisions are based on the evaluation of a risk calculated with this kind of model.

In their attempts to overcome or reduce model instability, some authors have suggested taking macro-economic factors responsible for this phenomenon into account (Mensah, 1984; Platt et al., 1994; Grice and Dugan, 2003; Pompe and Bilderbeek, 2005). They also showed that, by using some economic indicators (growth rate, interest rate, inflation rate, oil prices . . .) to weight traditional explanatory variables, it became possible to stabilize results. However, this solution is

applicable only *a posteriori* when one knows what the nature of the macro-economic changes was, and thus how to mitigate their effects. But, *a priori*, no one knows what should be done. Other authors demonstrated that one could take advantage of sampling variations caused by changes in the economic environment, and that one might improve model accuracy in the short term by using measures representing variation of ratios over time (standard deviation, coefficient of variation), but they did not study the stability of model accuracy in the long term (Dambolena and Khoury, 1980; Betts and Belhoul, 1987).

The latter approach implicitly acknowledges that the temporal stability of a model might be improved by using data collected over several consecutive years. This idea is close to the fact that history is a critical explanatory "variable" of business survival. Indeed, one knows that bankruptcy, in most cases, is the result of a long process (Laitinen, 1991), and that a firm's history strongly influences its ability to withstand failure. Thus, some companies can delay the onset of bankruptcy for many years, even though their financial profile shows that they should fail rapidly, whereas others manage to recover even though nothing suggests that an improvement may happen (Hambrick and D'Aveni, 1988). A firm's financial health measured at time t cannot be reduced to its situation measured at time $t - 1$ alone, although this is the underlying assumption of most failure models. However, this idea has been little explored so as to improve model accuracy (Balcaen and Ooghe, 2006), and even less to increase their stability over time

For this reason, we have decided to study this issue and to show how to improve stability by using data that measure the evolution of a company's financial health over several consecutive years, although not in the same way as previous research. Instead of using financial variables measured at different time intervals to design a model, we chose to use these variables to estimate "trajectories of corporate collapse", then to use these trajectories to make forecasts. We built on Laitinen's (1991) study, which assumed that the performance of a model heavily depends on its ability to account for the different paths companies may take during the years leading to their collapse. We also built on the hypothesis that there may be a finite number of trajectories, that these trajectories may be more stable over time than financial variables used to assess a company's performance, and hence that their results are likely to exhibit similar stability. We used different Kohonen maps to design trajectories. First, a map was calculated to delimit boundaries between areas representing various stages of company financial health. Secondly, we analysed how companies moved over time within these areas and we estimated a typology of behaviours we called "trajectories", using a set of Kohonen maps to quantize these behaviours. Third, we used this typology to forecast financial

5

failure and the performance achieved with trajectories was compared to that achieved using traditional failure models that were designed with discriminant analysis, logistic regression, survival analysis (Cox's method) and a neural network. As we looked at the influence of economic cycles on the stability of results, which appear to be the main factor of data non-stationarity, models were estimated and tested using data collected from different time periods (growth and downturn).

3. Samples and methods

Data were selected from a French database (Diane), which provides financial data on more than one million French firms. We only chose companies required by law to file their annual reports with the French commercial courts. We also chose companies in the same activity (retail) and of the same size (assets less than €750,000), to control for size and sector effects. We only selected income statement and balance sheet data, which have been the main sources of information for failure models since Altman (1968). This set of data was used to calculate ratios, and these ratios were subsequently used to design models. As we needed sufficient data to compute trajectories, we selected companies in operation for at least six years, keeping the same time frame as Laitinen (1991).

3.1. Covered periods

To collect data from different economic periods, we analysed changes in the French economic situation between 1991 and 2009. Over these 19 years, France experienced two recessionary periods and one downturn period[4]. The first recession occurred between March 1992 and June 1993, the downturn[5] occurred between March and December 2001, and the second recession began in March 2008. Figure 1 shows the changes in both gross domestic product (GDP) and business failure growth rates between 1991 and 2009, and clearly illustrates how downturns were preceded and followed by periods of growth, some more pronounced than others. We noticed a period of recovery and growth between 1993 and 2000, after the 1992 recession, despite a slight fall in GDP in 1996. In addition, the downturn which occurred in 2001 had an influence on the economy until late 2002, and growth slowly resumed in 2003 and continued to increase until early 2008.

---

[4]For economists, a recession occurs if GDP (gross domestic product) growth is negative for two or more consecutive quarters.

[5]The period between March and December 2001 can't be considered a recession since GDP growth was not negative for at least 2 quarters. GDP growth was extremely low over the second and third quarter, and became negative only over the last quarter (Source: macro-economic database – National Institute of Statistics and Economic Studies – http://www.bdm.insee.fr ; OECD.Statextracts – http://stats.oecd.org).

The different time periods during which we collected data were chosen based on Figure 1. We chose three periods; hence we designed three sets of models. The first was calculated with data collected from 1998 and was tested with data from 1999 and 2000 (estimation and test over a period of growth). The second set was calculated with data collected from 2000, and tested with data from 2001 and 2002 (estimation over a period of growth, and test over a period of downturn). Finally, the third set was designed using data collected from 2002, then tested with data from 2003 and 2004 (estimation over a period of downturn and test over a period of growth).

We wished to analyse model performance using data collected over a downturn period, either for estimation and test purposes. This would have required the downturn period to last at least three years: one year to collect data for estimation tasks, and the following two years to collect data for prediction tasks. Unfortunately, no recessionary or downturn periods that occurred in France between 1990 and 2011 have lasted this length of time. The recessions that occurred between 1992 and 1993 and between 2008 and 2009 are both immediately followed by a period of growth. And the situation is the same after the downturn that occurred between 2001 and 2002, with 2003 also being a period of growth.

3.2. Sample selection

We selected seven samples. Table 3 shows the years for which we collected data, the number of companies per sample and how samples were used for estimation or prediction tasks or both. For each sample collected at time t, company status (healthy vs. failed) was assessed at time $t+1$, with an average lag of 12 to 18 months. Balance sheets and income statements were selected from six consecutive years (t to $t-5$) and firms were chosen at random from among those in the database when they complied with the criteria described above.

3.3. Variable selection

Choosing a subset of variables from an initial set is essential to the parsimony of a model, but also essential for its accuracy and generalization ability. This task is difficult because the evaluation criterion used to select variables is often non-monotone. Indeed, the best subset of p variables rarely includes the best subset of q variables, where $q < p$. Faced with this non-monotonic character, only an exhaustive search of all possible combinations will lead to the best subset(s). But the resulting combinatorial explosion often makes these searches impossible. It is for that reason that most methods rely on heuristic procedures that carry out a limited search in the space of all combinations. These procedures are made up of three basic elements: a search method that

7

explores a subspace of all possible combinations and generates a set of candidate solutions; an evaluation criterion to evaluate the subset under examination and select the best one(s); a stopping criterion to decide when to stop the search method.

When the evaluation criterion is monotone, "complete" methods may find an optimal solution without evaluating all possible combinations, like Branch and Bound (Narendra and Fukunaga, 1977). Several variants derived from Branch and Bound can deal with situations where the criterion is non-monotone. With some of them, such as Approximate Monotonic Branch and Bound, the optimality of the solution is no longer guaranteed. However, with some others (Duarte Silva, 2001; Brusco and Steinley, 2011), the optimality can be reached but the size of the set of variables must be limited. These latter methods offer a valuable alternative to sequential methods, including stepwise methods, which usually only find local minima, but they work well only with a moderate number of variables (between 30 and 50) so that the computation times are not prohibitive.

Heuristic or "sequential" methods are used to relax the monotonic assumption that Branch and Bound imposes on the evaluation criterion and represent an alternative to complete methods, which are able to find a global minimum, when the number of variable is large. The simplest methods start with an empty set of variables, then add them one at a time (forward search); others start with all variables, then remove them, also one at a time (backward search). These methods produce results quickly, but they lead to non-optimal solutions because they search only a small part of the space. To increase the size of the search space, some methods such as Plus l – Take Away r alternate forward and backward steps (stepwise methods). Others called floating methods, derived from Plus l – Take Away r, also alternate forward and backward steps, but using a constraint on the evaluation criterion. Their advantage lies in dynamically determining the number of variables that are to be added or removed, as opposed to others where this number is fixed a priori (Jain and Zongker, 1997).

Finally, "random" methods reduce the risk of a method getting trapped in a local minimum. They start by choosing at random a set of variables, and then search using either a sequential strategy (Simulated Annealing) or a random one (genetic algorithms).

Once a search procedure has chosen a subset of variables, this subset must be evaluated. Some evaluation criteria rely solely on the intrinsic characteristics of the data without using the inductive algorithm (filter methods), that is to say the algorithm that will finally use the selected variables. Some others rely on the performance of the inductive algorithm when using the variables that are to be evaluated (wrapper methods). Independent criteria are either distance measures (e.g.,

8

Wilks' Lambda), information measures (entropy), dependence measures (mutual information) or consistency measures. When used in conjunction with some algorithms, these criteria do not always lead to optimal results, or they are difficult to use. Therefore, one uses the induction algorithm (Kohavi and John, 1997) for evaluation: each set of variables selected at any given time is then evaluated based on the generalization error of the model designed with these variables.

Without a suitable stopping criterion, a selection process could run until all possible combinations are evaluated. Many criteria may be used to interrupt a search. Most of the time these criteria take the form of a maximum number of iterations, a predefined number of variables, the absence of improvement after addition or removal of variables, or the achievement of optimal predictive ability. These criteria depend on computation heuristics and sometimes on statistical tests.

When it comes to choosing a variable selection method, one has to therefore choose a combination of three of the aforementioned techniques and criteria – a choice that is not necessarily straightforward. Indeed, depending on the selection method being considered, some techniques or criteria that have just been presented cannot be used, either because they are difficult to implement and cause intractable computational problems, or because they are ill-suited to the modelling technique. This is, for example, the case of evaluation criteria that rely on the hypothesis that input-output variable dependence is linear or that input variable redundancy is well measured by the linear correlation of these variables, and that are clearly not suited to neural networks (Leray and Gallinari, 1998). And even if the selection technique fits the modelling method, the question of the final choice and that of the role of the user in this choice always arises. This is the reason why some authors suggest that several criteria should be used simultaneously to select variables and that a certain degree of subjectivity might be involved in the selection process (Duarte Silva, 2001).

As a consequence, we used several selection methods and only variables that were most often selected were finally chosen to design our models. This approach corresponds to the idea proposed by (Murray, 1977) who thought that a good measure of the value of a variable for future classification would be the number of times it was selected by different techniques. The selection process was therefore organized as follows.

We calculated 41 financial ratios that can be broken up into a few, somewhat arbitrary categories that best describe company financial profiles: solvability, liquidity, profitability, financial structure, activity and rotation. We then selected, within this initial set of ratios, those that will be used to design models. We decided to choose a few subsets that are not sample- and selection-criteria dependent, but that best fit the three periods being examined. Therefore, we selected an initial

9

subset with data from 1998 (period of growth), a second subset with data from 2000 (period of growth) and a third with data from 2002 (period of downturn). For each period, we used the following procedure to select variables.

We chose six variable selection techniques commonly used with the modelling methods we chose for our experiments. We selected three techniques that are commonly used with discriminant analysis, logistic regression and Cox's method, and three others that are well suited for the neural network (multilayer Perceptron). We did not choose any method dedicated to Kohonen maps. Indeed, the way we used these maps is very different from the way they are traditionally used and therefore, there was no guarantee that a variable selection method tailored to Kohonen maps would have been, within the framework of our study, more useful than another.

The first technique we chose, which is tailored to discriminant analysis, relies on a stepwise search procedure, a Fisher F test as a stopping criterion, and a Wilks' Lambda as an evaluation criterion. The second and third methods, both tailored to logistic regression and Cox's model, respectively rely on a forward stepwise search and a backward stepwise search, with a likelihood statistic as an evaluation criterion of the solutions and a Chi2 as a stopping criterion in both cases. The last three techniques are well suited for the neural network. The fourth is a zero-order technique which uses the evaluation criterion designed by Yacoub and Bennani (1997). The fifth technique is a first-order method that uses the first derivatives of network parameters with respect to variables as an evaluation criterion. Finally, the sixth relies on the evaluation of an out-of-sample error calculated with the neural network. With these three criteria, we divided the learning sample into two parts: half of the sample was selected at random and used to select variables, while the other half estimated the generalization error of each subset. We used only a backward search rather than a forward or a stepwise search. As the search procedure involved successive removal of variables, the network was retrained after each removal, and the selection procedure was performed until all variables were removed (Leray and Gallinari, 1998). The results of a selection with a neural network strongly depend on the initialization of its weights. This is why we repeated the selection process with each criterion 10 times and finally, the set of variables that led to the lowest error was chosen. From a general standpoint, when a network is used during a selection process, its parameters are determined *a priori* (Leray and Gallinari, 1998). This solution does not necessarily lead to the best architecture, but it is faster than determining the parameters during selection. We then decided to estimate network parameters beforehand. To compute these parameters, we ran a set of experiments. At random, we drew 50 sets of variables from among those first selected. For

each set of variables, we tested several combinations of parameters: learning steps (from 0.05 to 0.5, with a 0.05 step) and the number of hidden nodes (from 2 to 15). We used only one hidden layer. All these figures were derived from those traditionally used with an MLP in the literature. For each combination of parameters, the error was estimated using a 10-cross-validation technique and the architecture that led to the lowest error was chosen. All networks used in this study were made up of one hidden layer, one output node and an activation function in the form of a hyperbolic tangent. We used the Levenberg-Marquardt algorithm, as an optimization technique during the learning process.

When all selections were done with each method, variables that were selected at least twice were chosen to design our models, but highly correlated variables were removed because correlation leads to model instability (Mensah, 1984). When the correlation between two variables was greater than 0.7, one of them was removed. We chose 0.7 as Atiya (2001) and Leshno and Spector (1996) did in the same context as that of this study. When deciding on which variables to remove, we used the following procedure. When one of the two variables was likely to give too much weight to a given financial dimension among those that were represented in a set of variables, it was discarded. The financial dimensions that are most often captured by variables used to design bankruptcy models are liquidity, solvency, profitability and financial structure, as suggested by many authors since Gupta (1969). We then managed to balance the weight of each main dimension captured by each set of variables. But when neither of the two variables was likely to overweight any given dimension, the variable that appeared to be the less "relevant" in the financial literature, given the issue we studied, was removed. The fact that a variable belonged to a certain financial dimension was assessed using the financial literature and confirmed using a principal component factor analysis.

3.4. Modelling methods

Several modelling methods were used to design prediction models: a Kohonen map, to calculate trajectories, and three traditional methods such as discriminant analysis, logistic regression and a neural network.

Models designed with discriminant analysis, logistic regression and the neural network are traditionally mono-period models which rely on a snapshot of a company's financial profile taken at a particular point in time. However, models designed with a Kohonen map rely on data that measure changes to a firm's financial health over a number of consecutive years (i.e. six years). Then, to ensure the fairness of the comparison with the Kohonen map, and to control for the influence of

11

this difference between data used with each method (single period data vs. time-series data), we designed several multi-period models. We first used a survival analysis (Cox's method) which is specially designed to deal with time-series data. This method was used with data measured over six years. Secondly, we added models designed with discriminant analysis, logistic regression and the neural network that used data collected over different time periods. With these models, each explanatory variable was measured at different points in time. For example, a model that was designed with five financial ratios and that used data collected over six years is made up of thirty explanatory variables. We used two time periods (three and six years) to control for the influence of the number of variables on model accuracy, that is to say the influence of the number of parameters to be estimated. Indeed, in general, the more parameters to be estimated on a given data set, the greater the generalization ability of a model is difficult to achieve.

3.4.1. Discriminant analysis

Discriminant analysis is a classification method with the aim of classifying objects in two or several groups using of a set of variables. To design the classification rule, the method attempts to derive a linear combination of independent variables that will best discriminate between previously defined groups, which in our case are sound and unsound firms. This is achieved using a procedure of maximizing the between-group variance relative to the within-group variance. Discriminant analysis then computes a score z according to:

$$z = \sum_{i=1}^{n} (x_i w_i + c) \tag{1}$$

where w_i represents the discriminant weights, x_i indicates the independent variables (e.g., financial ratios) and c is a constant. Each firm is assigned a single discriminant score which is then compared to a cut-off value which determines the group the company belongs to.

Discriminant analysis exhibits optimal discrimination abilities when the joint distribution of independent variables is multivariate normal and when their variance-covariance matrices are equal (Wald, 1944). But, variables that are commonly used to design bankruptcy prediction models (financial ratios) rarely meet these requirements. However, the robustness of discriminant analysis against departures from these conditions for optimality makes it a widely used method.

12

3.4.2. Logistic regression

Logistic regression is often used in conjunction with or instead of discriminant analysis to relax the conditions for optimality that the latter method imposes on the data. This is particularly the case when variables used to design models, such as financial ratios, exhibit characteristics that depart severely from these conditions.

A logistic regression function computes a probability score z for each observation (firm) to be classified, where:

$$z = \frac{1}{1 + e^{-\sum_{i=1}^{n}(x_i w_i + c)}} \tag{2}$$

where x_i represents the independent variables and c is a constant. The coefficients w_i of the function are calculated using maximum likelihood estimation. As with a discriminant function, an observation will be classified in one of two groups depending on its score.

3.4.3. Neural network

Neural networks, like discriminant analysis or logistic regression, are commonly-used classification methods in the field of bankruptcy prediction. Unlike discriminant analysis and logistic regression, neural networks do not represent the relationship between the independent variables and the dependent variable with an equation. This relationship is expressed as a matrix containing values (also called weights) that represent the strength of connections between neurons. In this study, a multilayer Perceptron (MLP) was used to perform the classification task. From a general point of view, a MLP used for classification tasks with two groups is made up of three layers: an input layer with n neurons, one per explanatory variable; one hidden layer with m neurons; and an output layer with one neuron. The layers are linked together, and the relationships between neurons are represented by weights: the weights w_{ij}^1 represent the relationships between the neurons of the input layer (x_i) and the neurons of the hidden layer. The weights w_j^2 represent the relationships between the neurons of the hidden layer (h_j) and the output neuron.

If one considers a classification task of observations into two groups to be achieved by such network, the vector x represents the explanatory variables, and the output neuron represents the result of the classification: class 1 or class 2. The values go through the network as a result of the activation function of each neuron. The activation function transforms input into output. The input value of a hidden neuron (h_j) is the weighted sum of the input neurons $\sum_{i=1}^{n}(x_i w_{ij}^1)$ and its output is $f(\sum_{i=1}^{n} x_i w_{ij}^1)$. The output of the output neuron is $f(\sum_{j=1}^{m} h_j w_j^2)$. The transformation of the input is done through a squashing function f, most often a hyperbolic tangent or a logistic function. This

13

transformation allows the network to take into account the non-linearity that may exist between its input and its output. The weights of the network are estimated through a learning process. During this process, network weights are tuned to values that allow the network to achieve a good classification rate with data used during the learning phase, but also good prediction ability when using data that were not used during this phase. Once the learning process is done, the network can be used for forecasting tasks.

An MLP does not require distributional assumptions of the independent variables and is able to model all types of non-linear functions between the input and the output of a model. This universal approximation capability, assessed Hornik et al. (1990), and the ability to build parsimonious classification rules make them powerful models.

3.4.4. Cox's method

Cox's method is one of the techniques known as survival analysis which allows the time that will elapse before a particular event occurs, such as bankruptcy, to be taken into account. This method is completely different from the previous ones. With discriminant analysis, logistic regression or neural networks, bankruptcy prediction is achieved using a classification rule. With a Cox's model, bankruptcy prediction is done using a timeline over which a firm is characterized by a specific lifetime. Lifetime distributions in a given population can be represented by two functions: a survival function and a hazard function. The survival function $s(t)$ represents the probability that a firm will survive past a given time t, and the hazard function $h(t)$ represents the instantaneous rate of failure at a given time t. There are different ways of assessing the survival and the hazard functions that depend on the assumptions about the relationships between these functions and a set of explanatory variables. With Cox's method, this relationship can be represented as:

$$h(t) = h_0(t)e^{\sum_{i=1}^{n} x_i w_i} \qquad (3)$$

where $h_0(t)$ corresponds to the baseline hazards and describes how the hazard function changes over time and $e^{\sum_{i=1}^{n} x_i w_i}$ corresponds to the way the hazard function relates to explanatory variables x_i. The regression coefficients w_i are calculated with a method similar to the maximum likelihood method.

The survival function $s(t)$ of a given company can be defined as:

$$s(t) = e^{-h(t)} \qquad (4)$$

14

As with a discriminant or a logistic function, a firm will be classified in one of two groups depending on its survival function.

3.4.5. Kohonen maps

Kohonen maps were originally designed to deal with clustering issues. A map consists of a set of neurons organised on a square grid most of the time. Each neuron is represented by an n-dimensional weight vector $w = (w_1, \ldots, w_n)$, where n is the dimension of the input vectors (i.e. number of variables used to represent observations). The weights of a map are calculated through a learning process during which the neurons learn the underlying patterns within the data. During this process all data vectors are compared to all weight vectors through a distance measure. For each input vector, once the nearest neuron is found, its weights are adjusted so as to decrease the distance between the input vector and this neuron. The weights of all neurons located in its neighbourhood are then also adjusted, but the magnitude of the variation is proportional to the distance between them on the map. Throughout the learning phase, the neighbourhood radius gradually shrinks, depending on a function to be defined *a priori*. This procedure is repeated until a stopping criterion is reached.

When the learning process is done, the resulting map is a non-linear projection of an n-dimension input space onto a two-dimensional space, which preserves the structure and topology of input data relatively well (Cottrell and Rousset, 1997): two companies that are close to each other in the input space will be close on the map. As the classes are known (failures vs. survivors), each neuron can be labelled with the label of the class for which it appears as a prototype. To do so, all input vectors are once again compared to all neurons. The percentage of companies in each class that are the closest to each neuron is then computed. Finally, the neurons are labelled with the label of the class whose percentage is the highest. The algorithm used during the learning phase of the map can be described as follows:

- Step 1: set the size of the map, using l lines and c columns, then randomly initialize the weights;

- Step 2: set the input neuron values $x = (x_1, \ldots, x_n)$ using data from one company;

- Step 3: compute the distance between vector $(x_1, \ldots, x_n)$ and the weight vector $(x_{k1}, \ldots, w_{kn})$ of each neuron w_k and select neuron w_c with the minimum distance:

$$\|x - w_c\| = min\|x - w_k\|$$

- Step 4: update weights within the neighbourhood of w_c:

  $$w_k(t+1) = w_k(t) + \alpha(t)h_{ck}(t)[x(t) - w_k(t)]$$

  where t is time, $\alpha(t)$ the learning step, $h_{ck}(t)$ the neighbourhood function, and $x(t)$ the input vector. The neighbourhood function is traditionally a decreasing function of both time and the distance between any neuron w_k on the map and neuron w_c that is the closest to the input vector at time t.

- Step 5: repeat step 2 to step 5 until t reaches its final value.

3.4.6. Kohonen map and trajectory design

A trajectory represents the variation of company's financial health over time. It also represents the way a firm moves in at-risk regions, over a number of consecutive years. A trajectory is then a sequence of positions within a space at risk, over a given period. We used a Kohonen map to design the space at risk, and a few others to design trajectories.

A map made up of 100 neurons was used to design the space at risk – 10 per row and 10 per column[6]. Three different spaces were then designed: one with data from 1998; one with data from 2000; and a last one with data from 2002. We used Sammon's mapping method (Sammon, 1969) to make sure that the input space was well approximated with a two-dimensional map and that no significant folding or stretching was visible on the map

Once the design of each map was completed,[7], we looked for neurons that can be considered prototypes of healthy and failed firms. To do so, we compared data used to design the maps and all neurons. We then calculated the percentage of sound and unsound firms that were closest to each neuron. Finally, neurons were labelled with the label of the class (healthy or failed) whose percentage was higher. When neurons are labelled, the map can be used to visualize the location of companies belonging to each class. It gives a complete picture of the proximity between failed and non-failed firms on the map, and makes it possible to represent a "failure" and a "non-failure space" and the boundaries between them.

Once the map was designed, company trajectories were computed: a trajectory is a path along which a company moves on the map from one neuron to another (i.e. from one region at risk to another) over a six-year period. These at-risk regions can be considered the hierarchies of financial profiles that best summarize all company financial situations. As we have collected data over

---

[6]This figure is somewhat arbitrary but it corresponds to usual empirical practices (Cottrell and Rousset, 1997).

[7]See Cottrell et al. (1998) regarding the algorithm and the procedure used to design a map.

16

six-year periods, each company can be represented using six vectors, one for each year. To locate the position of a company on the map, we computed the distance between all neurons and the six vectors. The neurons that are the closest to each vector then represent the different positions of a company on the map over time. Each sequence of six positions can be considered a trajectory.

However, as a map is made up of 100 neurons, it becomes impossible to analyse and visualize all possible trajectories. To reduce the number of combinations, we attempted to group neurons into a few super-classes. Each class of neurons was analysed separately to look for groups only representing healthy companies, and other groups only representing unhealthy companies. Neurons were then grouped into a small number of groups called super-classes[8] using a hierarchical ascending classification[9].

We then ranked the super-classes by the financial health of the companies they represented, ranging from companies in very good shape (super-class 1) to those in very bad shape (super-class n, n>1). Financial ratios were used to establish the hierarchy. Once established, we computed the different trajectories in keeping with the initial position of each company on the map over the first year of each period studied (that is, the positions in 1993, 1995 and 1997). We first calculated company trajectories whose initial positions in 1993, 1995 and 1997 were super-class 1, then company trajectories whose initial positions were super-class 2, etc. There are as many sets of trajectories as super-classes on the map.

Figure 2 depicts three individual trajectories on the map designed with data from 2002. This particular map is made up of 6 super-classes. Super-classes 1, 2, 3 and 4 (healthy zone on the map) represent neurons which encode healthy companies, and super-classes 5 and 6 (bankruptcy zone) represent neurons which encode failed companies. On this map, each set of lines, depicted with a different colour, shows the behaviour of a company. The steps are numbered and each one encodes a position on the map within a year – 1 is the position in 1997, 2 in 1998, 3 in 1999, 4 in 2000, 5 in 2001 and 6 in 2002.

The first trajectory (black lines) exhibits the behaviour of a company (firm 1) that stayed healthy for six years and whose initial (1997) and final (2002) positions on the map was super-class 1. The

---

[8]We analysed a few partitions made up of six to eleven super-classes, and we finally chose the best partition in terms of homogeneity. The homogeneity was assessed using the three best homogeneity indexes mentioned in the research carried out by Milligan (1981).

[9]We used three aggregation criteria (average linkage, complete linkage and Ward criterion) and each neuron was assigned to a distinct super-class, and then labelled. A neuron was labelled with the class predicted using these two or three criteria. When the three criteria led to different results, a neuron was labelled using a majority voting scheme, depending on the class of its nearest neighbours.

17

second one (gray lines) shows how a company (firm 2) moved slowly along a path to failure from an initial position in super-class 1 to a final position in class 6. In 1997, its situation was fairly good, but as time went by, its financial ratios progressively deteriorated and, finally, it went bankrupt in 2003. The third trajectory (white lines) is rather erratic. This firm (firm 3) was in bad shape in 1997 (initial position in class 5), and managed to recover two years later (position in class 1), but this remission was short. From 2000 to 2001 its situation worsened (position in class 5), only to get better in 2002 (position in class 1).

For each period, when all individual trajectories (one per company) were calculated, we then grouped these trajectories into prototype trajectories – one per super-class. For each super-class, we used a single-layer, six-neuron Kohonen map to compute the prototype trajectories. Six neurons were enough to correctly quantify all data, because, with more than six, some trajectories became indistinguishable from others, and with fewer, some no longer existed. With each learning sample (i.e. data from 1993 to 1998, 1995 to 2000 and 1997 to 2002), all prototype trajectories were labelled with the label of a class (sound or unsound) depending on a cut-off value described in the next paragraph below. Finally, we grouped all six-neuron maps into a final set (one per period), and we used it to complete the forecasts.

Figure 3 shows the distribution of prototype trajectories designed with the map depicted on Figure 2 over the period from 1997 to 2002. The six lines on Figure 3 display trajectories whose origin is super-class 1, 2..., 6 respectively on the map presented on Figure 2. On each graph, the scale of the X-axis corresponds to the six years, and the scale of the Y-axis corresponds to the six super-classes. The percentages shown in the columns are the proportion of companies belonging to each set of trajectories; the percentages shown below each graph represent the same proportion, but within each trajectory.

The first line displays the behaviour of companies belonging to super-class 1, that is, firms with the best financial health. The first four trajectories show that most of these firms never shifted to the "bankruptcy space", unlike the last two trajectories, which show that some of them ultimately went bankrupt.

Conversely, the last line displays, on the first two trajectories, how companies that were in bad financial shape in 1997 managed to improve, and on the last four trajectories it shows how other companies, also in bad shape, finally collapsed.

3.4.7. Cut-off value determination

The cut-off value used to discriminate between sound and unsound firms was calculated with two different methods. A set of forecasts were then made with each method.

Firstly, the cut-off value was estimated so as to maximize the overall rate of correct classifications. This is the most commonly used method in the bankruptcy literature. However, as previously mentioned, since the cost of misclassification between two models that exhibit equal performance is not symmetric, the best option is certainly the model that is able to minimize the type-I error (bankrupt firms predicted as healthy). This is the reason why many authors recommend to taking this cost into account while assessing the optimal cut-off value, and explicitly using it during the computation of a classification function. A second way of assessing the cut-off value was then used to take the observed expected cost of misclassification – also called resubstitution risk – into account (Frydman et al., 1985). Thus, the following objective function was minimized to estimate the cut-off value:

$$\text{Expected cost} = c_1 p_1 \frac{e_1}{N_1} + c_2 p_2 \frac{e_2}{N_2} \tag{5}$$

where c_1 and c_2 are the respective costs of type-I and type-II errors; p_1 and p_2 are the respective prior probabilities of bankruptcy and non-bankruptcy; e_1 and e_2 are the respective type-I (failed firm predicted as healthy) and type-II (healthy firm predicted as failed) errors; and N_1 and N_2 are the respective numbers of failed and healthy firms in the sample.

The main difficulty lies in specifying the values to be used with c_1 and c_2. Indeed, as suggested by Pacey and Pham (1990), c_1 and c_2 differ from firm to firm, but also from the situation of the user of the model and therefore on its own cost-of-error function. We then used the misclassification costs used by Frydman et al. (1985), the aim of whose study was also to compare different bankruptcy prediction models, where the cost of misclassification of healthy firms (i.e., c_2) is kept to 1, while the costs of misclassification of unsound firms (i.e., c_1) are respectively set to 1, 10, 20, 30, 40, 50, 60 and 70. As far as the prior probability of bankruptcy is concerned, we used the average probability of French firms belonging to the retail sector over the period studied – that is, 2%. This parameter was also used by Frydman et al. (1985) and by Tam and Kiang (1992).

With discriminant analysis, logistic regression and Cox's model, the cut-off value was estimated using the previously presented cost function, and in accordance with what has been done in the literature (Frydman et al., 1985; Tam and Kiang, 1992; Hopwood et al., 1994; Lee and Urrutia, 1996). With the neural network, the cost function to be minimized during the learning process

was changed so as to take into account the prior probabilities and the costs of misclassification (Tam and Kiang, 1992). Finally, with the Kohonen map, the cost function was used to label the trajectories that were used to make forecasts, in the same way as terminal nodes of a classification tree are labelled when using such a function (Frydman et al., 1985). Trajectories were labelled with the label of the class (healthy vs. failed) that minimized the observed cost of misclassification.

Consider a trajectory t that has to classify $n_i(t)$ objects from group i, and let N_i be the size of group i in the sample, with $i = 1, 2$. The risk of labeling trajectory t with the label of group 1 is defined as:

$$r_1(t) = c_2 p(2, t) = c_2 p_2 p(t|2) = c_2 p_2 \frac{n_2(t)}{N_2} \tag{6}$$

where $p(2, t)$ is the probability that a firm belongs to group 2 and is close to trajectory t and $p(t|2) = \frac{n_2(t)}{N_2}$ is the conditional probability of a group 2 firm being closed to trajectory t.

In a similar way:

$$r_2(t) = c_1 p_1 \frac{n_1(t)}{N_1} \tag{7}$$

As a consequence, a trajectory is labelled with the label of a class corresponding to the minimum risk.

3.4.8. Benchmarking scheme

Trajectory performance was benchmarked against that of models designed with traditional methods. For each period, we designed 10 models.

Discriminant analysis, logistic regression and the neural network were used with different sets of data. Over the first period we analysed, we used data from 1998 to calculate one-year period models. We then used data from 1996 to 1998 to compute three-year period models, and data from 1993 to 1998 to compute six-year period models. Finally, Cox's method and data from 1993 to 1998 were used to compute the last model.

The same scheme was applied for the two other periods studied: we used data from 2000 (2002) to calculate mono-period models, data from 1998 to 2000 (2000 to 2002) to compute three-year period models, data from 1995 to 2000 (1997 to 2002) to calculate six-year period models, and data from 1995 to 2000 (1997 to 2002) to calculate Cox's models.

3.5. Evaluation of model forecasting ability

The forecasting ability of one-year period models designed with discriminant analysis, logistic regression and the neural network, and using data from 1998, was assessed with data from 1999

20

and 2000. The forecasts were performed by comparing the predicted class achieved with a given model with the status of each company (healthy or failing). The same procedure was applied to the two other periods: the forecasting ability of models designed with data from 2000 and 2002 was assessed using data from 2001 and 2002, as well as data from 2003 and 2004. The forecasting ability of three- and six-year period models designed with discriminant analysis, logistic regression and the neural network, was estimated in a similar way, but using data from three and six consecutive years, respectively. Finally, with Cox's models, we also used the same procedure, with data from six consecutive years, for each period.

As far as trajectories are concerned, their performance was estimated as follows. We first computed the positions of companies on a map over a six-year period. The map, designed using data from 1998, was used to calculate trajectories with data from the periods of 1994 to 1999 and 1995 to 2000. We used the same procedure with the two other maps: the first was designed with data from 2000 and used to estimate trajectories with data from the periods of 1996 to 2001 and 1997 to 2002; the second map was designed with data from 2002 and used to estimate trajectories with data from the periods of 1998 to 2003 and 1999 to 2004.

For each period, forecasting was done by comparing all company trajectories with all prototype trajectories, using an Euclidean distance. A company was classified as healthy (or failed) over a given period if the prototype trajectory closest to its own trajectory was labelled as healthy (or failed). Table 4 describes how the different samples were used with each method to design and test all models.

To compute the generalization error of each model, we first calculated the predicted class of each company. With discriminant analysis, logistic regression, Cox's model and the neural network, the predicted class was assessed as follows:

$$
y_i^* = \left\{ \begin{array}{l} 0 \text{ (healthy) if score } y_i^{'} \text{ of company } i > y^* \\ 1 \text{ (failed) if score } y_i^{'} \text{ of company } i \leq y^* \end{array} \right\} \tag{8}
$$

where y_i^* is the predicted class of company i, $y_i^{'}$ is the score of company i and y^* are the cut-off values used to determine the boundary between the two classes.

With the trajectories, the predicted class was estimated with the following procedure. Among all t_n trajectories that were calculated, one seeks for the trajectory T which is the closest to the

21

trajectory of firm i as follows:

$$T(i) = \arg\min_n d(t_i, t_n) \tag{9}$$

where d is an Euclidean distance.

The predicted class y_i^* of firm i is then the class assigned to trajectory T. We estimated the classification error of each company as follows:

$$e_i = \left\{ \begin{array}{l} 1 \text{ if } y_i^* \neq y_i \\ 0 \text{ if } y_i^* = y_i \end{array} \right\} \tag{10}$$

were e_i is the classification error of company i, y_i^* is the predicted class of company i and y_i is the current class of company i.

Finally, we assessed the global classification error, type-I (misclassifying a failed firm) and type-II (misclassifying a healthy firm) errors of each model as follows:

$$\text{Global classification error} = \sum_{i=1}^{N} \frac{e_i}{N} \tag{11}$$

where e_i is the classification error of company i and N is the sample size.

$$\text{Type-I error} = \sum_{j=1}^{N_F} \frac{e_j}{N_F} \tag{12}$$

$$\text{Type-II error} = \sum_{k=1}^{N_H} \frac{e_k}{N_H} \tag{13}$$

where e_j is the classification error of failed company j, N_F is the number of failed firms, e_k is the classification error of healthy company k and N_H is the number of healthy firms.

4. Results and discussion

In the remainder of the paper, the expression "six-year period model" will solely refer to models designed with discriminant analysis, logistic regression and the neural network and data collected over six years, even if Cox's models and trajectories were also designed using six-year data. This expression is just used to differentiate models estimated with traditional methods over different time periods: one, three and six years.

Two different sets of estimations were calculated to assess model predictive ability. The first one

22

was calculated using a cut-off value that maximizes the overall rate of correct classifications. The second was assessed using different misclassification costs.

The first results presented below correspond to results computed without taking different misclassification costs into account. Table 5 shows the correct classification rates achieved with all models, and data from the three periods studied. When models are estimated and tested with data collected over periods of growth, their accuracy is quite similar, except for six-year period models designed with discriminant analysis, logistic regression and the neural network. But whatever their accuracy, their performances are particularly stable over time. Thus, for 1999 and 2000, trajectories lead to figures for correct classification of 81.2% and 81.9%, respectively. With data collected over one-year periods, logistic regression achieves an accuracy of 81% and 81.6%, respectively, compared to 81.3% and 81.3% for the neural network, 81.1% and 81.1% for Cox's method, and 80.2% and 81.2% for discriminant analysis. Model accuracy calculated with data collected over three-year periods is slightly lower than that estimated with data collected over one-year periods: logistic regression achieves an accuracy of 81.1% and 81.1% for 1999 and 2000 respectively, compared to 81.2% and 81% for the neural network and 80.7% and 80.9% for discriminant analysis. Finally, models designed with six-year period data lead to figures for correct classification far lower than the previous ones, and especially the model designed with the neural network: logistic regression leads to figures of 79.9% and 79.2%, compared to 79.8% and 78.8% for discriminant analysis and 78.6% and 76.4% for the neural network. These differences are certainly the result of model overfitting since the number of parameters to be estimated with six-year period models is much larger than the number of parameters used by one-year period models. It is interesting to notice that one-year and three-year period models behave in a similar way, and that there is no overfitting for the latter. By contrast, models designed over a period of growth lead to much poorer results when applied to data collected over a downturn. The accuracy of models estimated with data from 2000 falls particularly when we use data from 2002 – that is, data that characterized companies when the effects of the 2001 downturn were fully felt. Cox's method achieves a correct classification rate of 78.8%, as opposed to 78.7% for trajectories, 76.3% and 76.6% for the neural network using one- and three-year period data, respectively, 74.7% and 75.1% for discriminant analysis using one- and three-year period data, 74.3% and 74.8% for logistic regression using one- and three-year period data. Model accuracy continues to fall with six-year period models: discriminant analysis leads to an accuracy of 73.4%, compared to 72.2% for logistic regression and 72% for the neural network. Finally, when models are estimated with data collected over a period of downturn and used for

23

forecasts with data collected over a period of growth, results are not significantly better. With data from 2004, correct classification rates range from 71.4% for the neural network (six-year model) to 80% for trajectories.

On the whole, discriminant analysis, logistic regression and the neural network are not able to achieve better results than Cox's method when using multi-period data. Moreover, trajectories and Cox's method are more resistant than others to changes in economic conditions. They also do significantly better than discriminant analysis and logistic regression, and to a lesser extent, than the neural network. In that sense, they provide a clear gain in stability. To ascertain this fact, we compared the results achieved with all methods and samples from 2000, 2002 and 2004. For this purpose, we carried out a test for differences between the different correct classification rates. The significance levels of this test are shown in Table 6.

Symbols in Table 6 clearly show that the differences between correct classification rates achieved with trajectories and Cox's method on one hand, and discriminant analysis and logistic regression, on the other, are almost all statistically significant (at the conventional threshold of 5%) when changes in economic conditions occur between the period during which models are estimated and during which they are used for forecasts. But when the p-values exceed this threshold of 5%, they always remain below 10%. Conversely, these differences are not all significant when correct classification rates achieved with trajectories and Cox's method are compared to those achieved with the neural network. If the comparison is done with six-year period models designed with the neural network, the differences are largely significant. If the same comparison is done with one- and three-year period models, there is no statistical difference with data from 2000 and 2002, whereas the differences are slightly significant with data from 2004 (p-values range from 5% to 20%).

We can speculate that the gap between the performances of the different models is the result of the fit between time-series data and the characteristics of the methods themselves since multi-period models designed with conventional methods are much more sensitive to changes in economic conditions than trajectories and Cox's model are. It is worth noting that the neural network performs relatively well with short time-period data (one and three years).

We then studied the aforementioned differences, but depending on whether companies are healthy or have failed. Table 7 shows the percentage of correct classifications for these two groups. This table indicates that all methods achieved rather stable forecasts when predicting the fate of sound firms. But when it comes to predicting the fate of unsound companies, results are much more unstable. When estimation and test data are collected over a period of growth, healthy companies

achieve correct classification rates ranging from 80.4% to 82.2% with data from 1999, as opposed to rates ranging from 80.3% to 84.7% with data from 2000. At the same time, failed companies achieve rates ranging from 75.6% to 81.9% and from 69% to 83.5% respectively. And the gap widens greatly with the collection of estimation data during a period of growth, as with the collection of test data during a downturn. In such circumstances, healthy companies achieve correct classification rates ranging from 79.3% to 83.7% with a sample from 2001 and from 77.3% to 86.7% with a sample from 2002. Conversely, for failed companies the figures range from 69.2% to 80.7% with a sample from 2001, and from 58% to 80.1% with a sample from 2002. Figures for classification rates achieved with data from the third period are between those of periods 1 and 2.

This result is consistent with the results of Pompe and Bilderbeek (2005), who noticed that any change in the economic environment makes it more difficult to predict the fate of failed firms than that of healthy firms. Indeed, they found that the moment in which a model's performance deteriorates coincided precisely with the moment in which an economic decline and a significant increase in the number of bankruptcies occur. One possible explanation for this phenomenon is that some firms that could have survived in a sound economic climate are no longer able to do so when the climate deteriorates. This reasoning also applies in the opposite situation, as shown in Table 7, with data collected over the third period. Indeed, a model that is designed during a period of downturn and used during a period of growth also fails to accurately forecast the fate of unsound companies.

Table 7 also shows how Cox's model, and especially trajectories, managed to do better than other methods at predicting the fate of failed firms. Besides, trajectories tend to predict the fate of unsound companies better than that of sound firms. Indeed, with a sample from 2002, correct classification rates of failed (healthy) companies achieved with trajectories are 80.1% (77.3%), compared to rates ranging from 58% to 73.5% (79.6% to 86.7%) for other models. Similarly, with a sample from 2004, figures are 80.1% (79.8%) for trajectories, as opposed to figures ranging from 60.5% to 77.7% (79.6% to 84.4%) for other models.

With multi-period models, the difference between correct classification rates of failed and healthy firms tends to increase compared to that of mono-period models, except with the neural network using three-year period data. In such circumstances, the neural network achieves a correct classification rate of failed companies that is on average lower than that achieved with Cox's model (and of course with trajectories), but far better than that achieved with all other models.

Differences between all classification rates were statistically assessed and the results of the test are

shown in Tables 8 and 9. These tables show that, in most cases, trajectories do significantly better than other models when predicting the fate of failed companies. It also shows that Cox's model behaves in a similar manner as trajectories, although it is not as accurate as trajectories, especially when it comes to failed firms. Finally, it points out that the neural network using three-year period data presents some similarities with Cox's model. Indeed, this model has the ability to better forecast the fate of failed firms than other models do, even if its global accuracy is weaker than that of Cox's model.

To deepen the results estimated with a cut-off value maximizing the overall classification rate, we analysed how models behave using different misclassification costs. As mentioned in section 3.4.7., the cost of misclassification of healthy firms (i.e., c_2) was kept to 1, while the costs of misclassification of unsound firms (i.e., c_1) were respectively set to 1, 10, 20, 30, 40, 50, 60 and 70. The prior probabilities of the failed and non-failed groups were fixed at 2% and 98%, respectively. The situation where the prior probability of bankruptcy is set to 2%, and the respective costs of misclassification of healthy and failed firms are set to 1 and 50, leads to classification results that are nearly equal to those obtained when the prior probability of bankruptcy is 50% and when the costs of misclassification are equal and set to 1.

For each period, we calculated the percentage of misclassified companies by model. In order not to overburden the presentation with excessive detail, only misclassification rates achieved with cost 1, 10, 30, 50 and 70 are presented. Tables 10, 11 and 12 show the percentage of misclassified companies by period and by model for different misclassification costs. These tables show that when the classification cost of failed firms is low, trajectories achieve an overall error rate lower than that achieved with all other models. As the cost increases, the errors achieved with all models tend to decrease faster than the error achieved with trajectories, but in nearly all situations this latter error remains the lowest. With data from 2000, all models (except Cox's model and trajectories) achieve an error ranging from 45.5% to 53% for cost 1, and from 16% to 23.5% for cost 70, whereas trajectories achieve an error of 39% for cost 1 and 13.5% for cost 70. With data from 2002 and all models (except Cox's model and trajectories), figures range from 41% to 45.6% for cost 1, and from 19.1% to 27.3% for cost 70; and with trajectories, figures are 29.2% for cost 1, and 17.5% for cost 70. With data from 2004, the situation is the same: all models (except Cox's model and trajectories) achieve an error ranging from 46.4% to 53% for cost 1, and from 22.3% to 26.5% for cost 70; and trajectories achieve an error of 33.3% for cost 1 and 19.7% for cost 70.

The results achieved with Cox's model are rather close to those achieved with trajectories, for all

costs and whatever the period. Only once, with data from 2002, when the cost is equal to 50, does Cox's model perform better than trajectories. In a few cases (data from 2000 with cost 10 and cost 50; data from 2002 with cost 70), the neural network using three-year period data achieves error rates lower than those obtained with Cox's model.

As far as misclassification rates of failed and non-failed firms are concerned, when the cost of misclassification of failed firms is low, all type-II errors are, on average, very low: with data from 2000, type-II errors range from 0% to 7.1%; with data from 2002, figures range from 1.8% to 10.3%; and with data from 2004, they range from 2.9% to 11.1%. By contrast, with data from 2000, trajectories and Cox's model achieve a type-I error of 76.3% and 81% respectively, compared to type-I errors ranging from 89.6% to 100% with the other models. With data from 2002, for trajectories and Cox's model figures are 67.5% and 83% respectively, and for other models they range from 94% to 100%. Finally, with data from 2004, figures for trajectories and Cox's model are 66.8% and 87%, respectively, otherwise they range from 93.0% to 100%.

When the cost increases, trajectories managed to achieve type-I errors that remain far lower than those obtained with other models in any situation. Cox's model also performed well in terms of type-I errors compared to discriminant analysis, logistic regression and the neural network, even if the neural network using three-year period data outperformed Cox's model in a few situations (data from 2000 and costs 10 and 50, data from 2002 and cost 70).

Based on the results used to calculate Tables 10, 11 and 12, we have assessed the resubstitution risks of each model for all costs of misclassification. Tables 13, 14 and 15 show the resubstitution risks of all models assessed with data from 2000, 2002 and 2004.

In these three tables, the comparison of trajectories with other models for costs ranging from 1 to 70 shows that trajectories slightly dominate Cox's model, and all other models substantially. In Table 13, trajectories dominate all models in terms of resubstitution risks for all but costs 1 and 10. In Table 14, aside from Cox's model for cost 50, trajectories also dominate all models. And in Table 15, trajectories continue to dominate all models. We must add that, for cost 20 that is not presented in Table 15, Cox's model and the neural network using three-year period data have a lower overall cost than that of trajectories. We can also observe that Cox's model, especially with data from 2002 and 2004, slightly dominates the neural network using one- and three-year period models.

All these tables show that a hierarchy exists between all models and that it can be broken up into three groups. The first group is made up of trajectories, Cox's model and the neural network using

27

one- and three-year period data. These models are more robust to any changes in the misclassification costs than the others. Then, the second group is made up of one- and three-year period models designed with discriminant analysis and logistic regression. The third and last group is made up of six-year period models.

The results we obtain with conventional modelling techniques of designing financial failure models are consistent with the results of many studies published in the literature. Firstly, the assumption that model prediction ability depends on data stationarity is validated: model accuracy is indeed dependent on the phase of the business cycle during which data collection occurs. Secondly, changes in economic conditions involve asymmetric effects on model accuracy depending on whether companies are healthy or if they have failed; forecasting the fate of failed companies becomes even more difficult as the magnitude of changes increases.

The results we obtain with trajectories show that, when company status (healthy vs. failed) is not taken into account, and if the economic environment remains stable, their performance is similar to that achieved with other models. However, if the economic environment changes, then trajectories achieved similar results to those achieved with Cox's model, but better than those achieved with other models. But, when company status is taken into account, model performance depends heavily on this status. When it comes to predicting the fate of unsound firms, trajectories are indeed less sensitive to economic changes than other models are. If changes are slight, trajectories perform as well as Cox's model or the neural network using three-year period data (see Table 7), but better than the other models. And if changes are great, type-I error (misclassifying a failed firm) is much lower than that achieved with all other models, and this result also holds when misclassification costs are taken into account – a point that is particularly important. With equal performance, a good model is a model able to minimize type-I error because, as we have noted, the cost of this error is far greater than that of a type-II error (misclassifying a healthy company). Moreover, when it comes to predicting the fate of healthy companies, trajectories perform slightly worse than other models.

5. Conclusion

In this study, we proposed a new way of assessing the evolution of company's financial situation over time using a Kohonen map to quantize this evolution into trajectories. We also proposed using these trajectories to do forecasting and we compared their prediction performance to that of discriminant analysis, logistic regression, Cox's method and a neural network, particularly focusing

28

on how to stabilize model accuracy over time.

Our results confirm that the accuracy of conventional models is closely related to changes that occur in firms' economic environment over the period during which models are estimated and used to do forecasting. Their accuracy is even worse when the growth differential between the model estimation period and the test period is large, whereas trajectories achieve much more stable results. Moreover, when model accuracy deteriorates, the error made with conventional models is caused mainly by a sharp fall in type-I error, whereas trajectories achieve rather well-balanced results for type-I and type-II errors.

Trajectories can therefore increase the reliability of failure models over time and reduce the cost of errors in decision-making that may occur in financial institutions. They can also help settle the issue about the time intervals beyond which a model must be re-estimated. However, the novelty of this method is also its weakness. The results obtained in this study require confirmation with other samples and other types of firms. But, beyond this issue, other uses of trajectories might be explored. For one thing, they can be used as a diagnostic tool, to assess a company's financial situation, as Sueyoshi and Goto (2009) suggested a revamped version of Altman's (1968) z-score for a similar use. Therefore, one should analyse their informational content with regard to that of other diagnostic methods. For another, they might be used to tackling conceptual issues – for example, how to better understand some financial behaviours that characterize firms, whether they are healthy or not. It would be valuable to study what this representation might bring to the understanding of the bankruptcy process itself.

Acknowledgment

References

Agarwall, V., Taffler, R. J., 2007. Twenty-five years of the Taffler z-score model: Does it really have predictive ability? Accounting and Business Research 37, 285–300.

Altman, E. I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance 23, 589–609.

Altman, E. I., Eisenbeis, R. A., 1978. Financial applications of discriminant analysis: A clarification. Journal of Financial and Quantitative Analysis 13, 185–195.

Altman, E. I., Lavallee, M. Y., 1981. Business failure classification in Canada. Journal of Business Administration 12, 147–164.

Atiya, A. F., 2001. Bankruptcy prediction for credit risk using neural networks: A survey and new results. IEEE Transactions on Neural Networks 12, 929–935.

Aziz, A., Emanuel, D. C., Lawson, G. C., 1988. Bankruptcy prediction: An investigation of cash flow based models. Journal of Management Studies 25, 419–437.

Balcaen, S., Ooghe, H., 2006. 35 years of studies on business failure: An overview of the classical statistical methodologies and their related problems. British Accounting Review 38, 63–93.

Barth, J. R., Brumbaugh, R. D., Sauerhaft, D., Wang, G. H., 1985. Thrift-institution failures: Causes and policy issues. In: Conference on bank structure and competition. Federal Reserve Bank, pp. 184–216.

Begley, J., Ming, J., Watts, S., 1996. Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. Review of Accounting Studies 1, 267–284.

Betts, J., Belhoul, D., 1987. The effectiveness of incorporating stability measures in company failure models. Journal of Business Finance and Accounting 14, 323–334.

Boritz, J. E., Kennedy, D. B., 1995. Effectiveness of neural network types for prediction of business failure. Expert Systems with Applications 9, 503–512.

Boritz, J. E., Kennedy, D. B., Sun, J. Y., 2007. Predicting business failure in Canada. Accounting Perspectives 6, 141–165.

Brusco, M. J., Steinley, D., 2011. Exact and approximate algorithms for variable selection in linear discriminant analysis. Computational Statistics and Data Analysis 55, 123–131.

Cayssials, J. L., Golfier, C., Kendaoui, L., 2009. Les défaillances en France : suivi mensuel, difficultés statistiques et impact économique. In: 7ème journée d'études sur les faillites. Université Paris Ouest - Nanterre.

Charitou, A., Neophytou, E., Charalambous, C., 2004. Predicting corporate failure: Empirical evidence for the UK. European Accounting Review 13, 465–497.

Coats, P. K., Fant, L. F., 1993. Recognizing financial distress patterns using a neural network tool. Financial Management 22, 142–154.

Cottrell, M., Fort, J. C., Pages, G., 1998. Theoretical aspects of the Som algorithm. Neurocomputing 21, 119–138.

Cottrell, M., Rousset, P., 1997. The Kohonen algorithm: A powerful tool for analysing and representing multidimensional quantitative and qualitative data. In: Mira, J., Moreno-Diaz, R., Cabestany, J. (Eds.), Lecture Notes in Computer Science. pp. 861–871.

Dambolena, I., Khoury, S., 1980. Ratio stability and corporate failure. Journal of Finance 33, 1017–1026.

Deakin, E. B., 1977. Business failure prediction: An empirical analysis. In: Altman, E. I., Sametz, A. W. (Eds.), Financial crises: institutions and markets in a fragile environment. John Wiley, New York, pp. 69–88.

Drezner, Z., Marcoulides, G. A., Stohs, M. H., 2001. Financial application of a Tabu search variable selection method. Journal of Applied Mathematics and Decision Sciences 5, 215–234.

Duarte Silva, A. P., 2001. Efficient variable screening for multivariate analysis. Journal of Multivariate Analysis 76, 35–62.

Elmer, P. J., Borowski, D. M., 1988. An expert system approach to financial analysis: The case of S and L bankruptcy. Financial Management 17, 66–76.

Fan, A., Palaniswami, M., 2000. Selecting bankruptcy predictors using a support vector machine approach. In: IEEE-INNS-ENNS International Joint Conference on Neural Networks. Vol. 6. Federal Reserve Bank, pp. 354–359.

Fanning, K. M., Cogger, K. O., 1994. A comparative analysis of artificial neural networks using financial distress prediction. International Journal of Intelligent Systems in Accounting, Finance, and Management 3, 241–252.

Finlay, S., 2011. Multiple classifier architectures and their application to credit risk assessment. European Journal of Operational Research 210, 368–378.

Frydman, H., Altman, E., Kao, D., 1985. Introducing recursive partitioning for financial classification: The case of financial distress. Journal of Finance 40, 269–291.

Gentry, J. A., Newbold, P., Whitford, D. T., 1985. Classifying bankrupt firms with funds flow components. Journal of Accounting Research 23, 146–160.

Gombola, M. J., Haskins, M. E., Ketz, J. E., Williams, D. D., 1987. Cash flow in bankruptcy prediction. Financial Management 16, 55–65.

Grice, J. S., Dugan, M. T., 2003. Reestimations of the Zmijewski and Ohlson bankruptcy prediction models. Advances in Accounting 20, 77–93.

Grice, J. S., Ingram, R. W., 2001. Tests of the generalizability of Altman's bankruptcy prediction model. Journal of Business Research 54, 53–61.

Gupta, M. C., 1969. The effect of size, growth, and industry on the financial structure of manufacturing companies. Journal of Finance 24, 517–529.

Gupta, Y. P., Rao, R. P., Bagchi, P. K., 1990. Linear goal programming as an alternative to multivariate discriminant analysis: A note. Journal of Business Finance and Accounting 17, 593–598.

Hambrick, D. C., D'Aveni, R. A., 1988. Large corporate failures as downward spirals. Administrative Science Quarterly 33, 1–23.

Hopwood, W., McKeown, J. C., Mutchler, J. F., 1994. A reexamination of auditor versus model accuracy within the context of the going-concern opinion decision. Contemporary Accounting Research 10, 409–431.

Hornik, K., Stinchombe, M., White, H., 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Networks 3, 409–431.

Jain, A. K., Zongker, D., 1997. Feature selection: Evaluation, application and small sample performance. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 153–158.

Kohavi, R., John, G. H., 1997. Wrappers for feature subset selection. Artificial Intelligence, Special Issue on Relevance 97, 273–324.

32

Lacher, R. C., Coats, P. K., Sharma, S. C., Fant, L. F., 1995. A neural network for classifying the financial health of a firm. European Journal of Operational Research 85, 53–65.

Laitinen, T., 1991. Financial ratios and different failure processes. Journal of Business Finance and Accounting 18, 649–673.

Lee, K., Booth, D., Alam, P., 2005. A comparison of supervised and unsupervised neural networks in predicting bankruptcy of korean firms. Expert Systems with Applications 29, 1–16.

Lee, S. H., Urrutia, J. L., 1996. Analysis and prediction of insolvency in the property-liability insurance industry: A comparison of logit and hazard models. Journal of Risk and Insurance 63, 121–130.

Leray, P., Gallinari, P., 1998. Feature selection with neural networks. Behaviormetrika 26, 145–166.

Leshno, M., Spector, Y., 1996. Neural network prediction analysis: The bankruptcy case. Neurocomputing 10, 125–147.

McKee, T. E., Greenstein, M., 2000. Predicting bankruptcy using recursive partitioning and a realistically proportioned data set. Journal of Forecasting 19, 219–230.

Mensah, Y. M., 1984. An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. Journal of Accounting Research 22, 380–395.

Milligan, G. W., 1981. A monte-carlo study of thirty internal criterion measures for cluster analysis. Psychometrika 46, 187–199.

Moriarity, S., 1979. Communicating financial information through multidimensional graphics. Journal of Accounting Research 17, 205–224.

Mossman, C. E., Bell, G. G., Swartz, L. M., Turtle, H., 1998. An empirical comparison of bankruptcy models. Financial Review 33, 35–53.

Moyer, R. C., 1977. Forecasting financial failure: A re-examination. Financial Management 6, 11–17.

Murray, G. D., 1977. A cautionary note on selection of variables in discriminant analysis. Applied Statistics 26, 246–250.

Narendra, P. M., Fukunaga, K., 1977. A branch and bound algorithm for feature subset selection. IEEE Transactions on Computer c-26, 917–922.

Odom, M. C., Sharda, R., 1990. A neural network model for bankruptcy prediction. In: the IEEE International Joint Conference on Neural Networks. pp. 163–168.

Ohlson, J. A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18, 109–131.

Pacey, J. W., Pham, T. M., 1990. The predictiveness of bankruptcy models: Methodological problems and evidence. Australian Journal of Management 15, 315–337.

Pinches, G. E., Mingo, K. A., Caruthers, J. K., 1973. The stability of financial patterns in industrial organizations. Journal of Finance 28, 389–396.

Platt, H. D., Platt, M. B., Pedersen, J. G., 1994. Bankruptcy discrimination with real variables. Journal of Business Finance and Accounting 21, 491–510.

Pompe, P. P. M., Bilderbeek, J., 2005. Bankruptcy prediction: The influence of the year prior to failure selected for model building and the effects in a period of economic decline. International Journal of Intelligent Systems in Accounting, Finance and Management 13, 95–112.

Richardson, F. M., Kane, G. D., Lobingier, P., 1998. The impact of recession on the prediction of corporate failure. Journal of Business Finance and Accounting 25, 167–186.

Sammon, J. W., 1969. A nonlinear mapping for data structure analysis. IEEE Transactions on Computers c-18, 401–409.

Sueyoshi, T., Goto, M., 2009. Methodological comparison between DEA (data envelopment analysis) and DEA-DA (discriminant analysis) from the perspective of bankruptcy assessment. European Journal of Operational Research 199, 561–575.

Taffler, R., 1983. The assessment of company solvency and performance using a statistical model. Accounting and Business Research 13, 295–307.

Tam, K. Y., Kiang, M. Y., 1992. Managerial applications of neural networks: The case of bank failure predictions. Journal of Risk and Insurance 38, 926–947.

Wald, A., 1944. On a statistical problem arising in the classification of an individual into one of two groups. Annals of Mathematical Statistics 15, 145–162.

Wilcox, J. W., 1973. A prediction of business failure using accounting data. Journal of Accounting Research 11, 163–179.

Wu, Y., Gaunt, C., Gray, S., 2010. A comparison of alternative bankruptcy prediction models. Journal of Contemporary Accounting and Economics 6, 34–45.

Yacoub, M., Bennani, Y., 1997. A heuristic for variable selection in multilayer artificial neural network classifier. In: International Conference on Artificial Neural Networks and Intelligent Engineering. Vol. 7. pp. 527–532.

Yang, Z. R., Platt, M. B., Platt, H. D., 1999. Probabilistic neural networks in bankruptcy prediction. Journal of Business Research 44, 67–74.

Zavgren, C. V., 1983. The prediction of corporate failure: The state of the art. Journal of Accounting Literature 2, 1–38.

Zmijewski, M. E., 1984. Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research 22, 59–82.
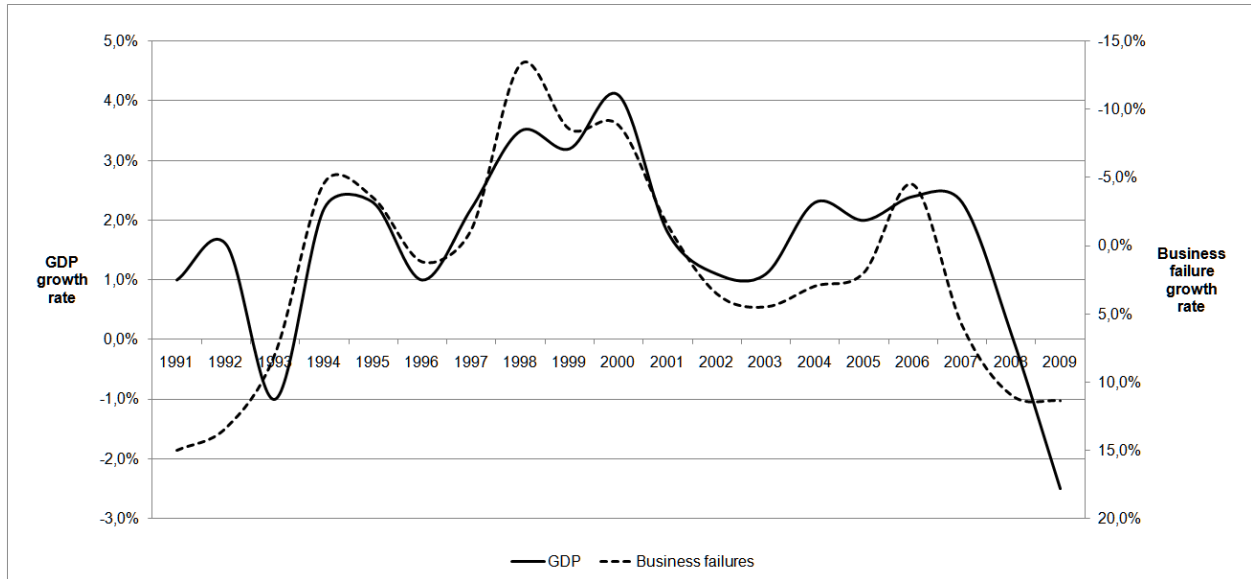
6. Figures

Figure 1: Changes in GDP and business failure growth rates in France between 1991 and 2009 – Source: Cayssials et al. (2009)
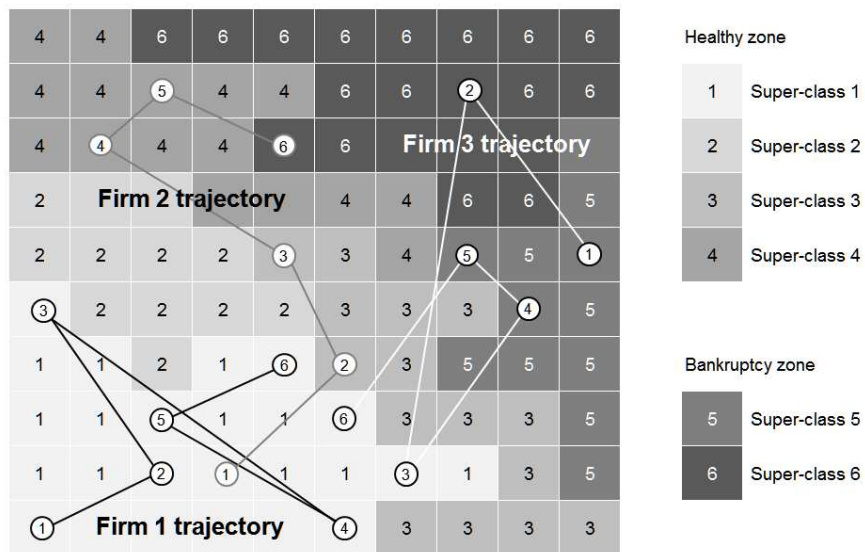
Figure 2: Examples of individual trajectories on the map designed with data from 2002
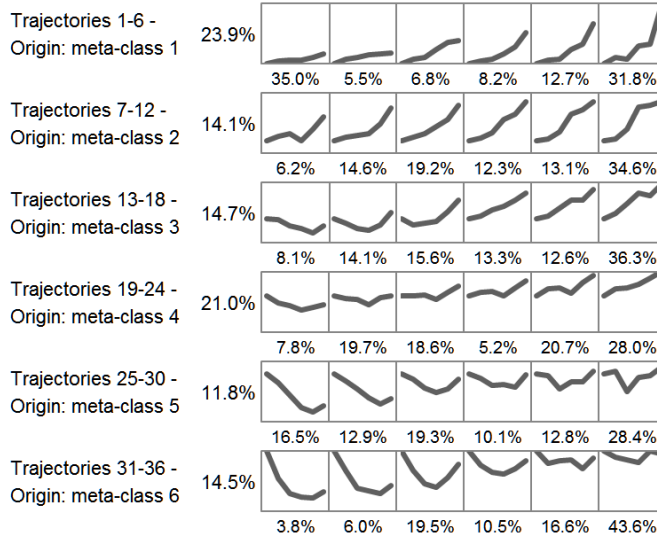
Figure 3: Distribution of prototype trajectories by initial company position on the map designed with data from 2002

7. Tables

Table 1: Results of the main financial failure models whose generalization ability was assessed in other studies

Models	Model accuracy			Sample size		Methods	Sectors	Years
	H	F	Total	H	F			
Altman (1968)	97.0%	93.9%	95.5%	33	33	DA	M	1946−1965
Altman and Lavallee (1981)	81,5%	85,2%	83,3%	27	27	DA	VS	1970−1979
Aziz et al. (1988)	97,9%	85,7%	91,8%	49	49	LR	VS	1971−1982
Barth et al. (1985)	66,0%	84,0%	74,9%	606	588	LR	F	1981−1983
Deakin (1977)	98,0%	89,0%	94,4%	86	57	DA	VS	1964−1971
Gentry et al. (1985)	87,9%	78,8%	83,3%	33	33	LR	VS	1970−1981
Gombola et al. (1987)			89,0%	244	77	DA	VS	1970−1982
Hopwood et al. (1994)	80,0%	73,0%	76,1%	30	37	LR	VS	1974−1985
Ohlson (1980)			96,1%	2058	105	LR	I	1970−1976
Platt et al. (1994)	95,5%	94,3%	95,2%	89	35	LR	GO	1982−1988
Taffler (1983)	95,7%	100,0%	97,8%	46	46	DA	M	1969−1976
Wilcox (1973)			93,8%	32	32	GRM	I	1954−1971
Zmijewski (1984)	99,5%	62,5%	97,7%	800	41	PR	I	1972−1978

H: Healthy - F: Failed.
DA: Discriminant analysis - GRM: Gambler's ruin model - LR: Logistic regression - PR: Probit.
F: Finance - GO: Gas and oil - I: Industry - M: Manufacturing - VS: Various sectors.
Figures presented in this table correspond to the best results when many results were computed.
Empty cells correspond to results that were not mentioned.

Table 2: Results of the main studies that have assessed the generalization ability of the models presented in Table 1

Models	Model accuracy			Sample size		Model types	Sectors	Years
Studies	H	F	Total	H	F			
Altman (1968)								
Fan and Palaniswami (2000)			64.3%	88	86	R	I	
Boritz and Kennedy (1995)	71.6%	65.6%	71.3%	56	56	O		1971−1984
Boritz et al. (2007)	65.6%	78.1%	71.9%	64	64	O		1987−2002
Boritz et al. (2007)	83.3%	41.7%	62.5%	24	24	R		1987−2002
Coats and Fant (1993)	100.0%	68.3%	87.9%	94	47	R	I	1970−1989
Lacher et al. (1995)	95.7%	83.0%	89.4%	47	47	O	I	1970−1989
Lee et al. (2005)			73.8%	42	42	R	I	1995−1998
Odom and Sharda (1990)			74.3%	64	65	R	I	1975−1982
Drezner et al. (2001)	37.8%	83.2%	60.5%	185	185	R	VS	1978−1997
Grice and Ingram (2001)	55.5%	70.9%	57.8%	831	148	O	I	1988−1991
Grice and Ingram (2001)	93.8%	54.7%	88.1%	831	148	R	I	1988−1991
Gupta et al. (1990)	90.0%	80.0%	85.0%	10	10	R	I	1971−1986
Moyer (1977)	88.0%	60.9%	75.0%	25	23	O	I	1965−1975
Moyer (1977)	81.8%	95.0%	88.1%	22	20	R	I	1965−1975
Begley et al. (1996)	74.9%	81.5%	75.2%	1300	65	O	I	1980−1989
Begley et al. (1996)	78.4%	78.5%	78.4%	1300	65	R	I	1980−1989
Moriarity (1979)	46.7%	57.1%	50.0%	15	7	O	I	1974−1975
Wu et al. (2010)	76.4%	45.4%	75.9%	49724	887	R	I	1980−2006
Altman and Lavallee (1981)								
Boritz et al. (2007)	51.6%	87.5%	69.5%	64	64	O		1987−2002
Boritz et al. (2007)	62.5%	58.3%	60.4%	24	24	R		1987−2002
Aziz et al. (1988)								
Mossman et al. (1998)	68.9%	71.1%	70.0%	45	45	R	VS	1980−1991
Barth et al. (1985)								
Elmer and Borowski (1988)	91.4%	86.7%	88.3%	60	60	R	F	1986
Hopwood et al. (1994)								
Richardson et al. (1998)	39.8%	96.9%	43.2%	2000	128	R	VS	1968−1990
McKee and Greenstein (2000)	94.0%	79.0%	93.9%	2468	19	R	VS	1986−1990
Ohlson (1980)								
Boritz and Kennedy (1995)	61.4%	91.1%	62.8%	56	56	O		1971−1984
Boritz et al. (2007)	62.5%	85.9%	74.2%	64	64	O		1987−2002
Boritz et al. (2007)	79.2%	58.3%	74.0%	24	24	R		1987−2002
Grice and Dugan (2003)	93.7%	59.1%	88.6%	889	154	R	I	1988−1991
Begley et al. (1996)	74.5%	89.2%	75.2%	1300	65	O	I	1980−1989
Begley et al. (1996)	85.1%	70.8%	84.4%	1300	65	R	I	1980−1989
Wu et al. (2010)	85.2%	81.3%	85.1%	49724	887	R	I	1980−2006
Platt et al. (1994)								
Yang et al. (1999)	86.7%	87.5%	86.8%	30	8	R	GO	1982−1988
Taffler (1983)								
Agarwall and Taffler (2007)	74.0%	96.0%	74.2%	27011	232	O	VS	1979−2003
Wilcox (1973)								
Fanning and Cogger (1994)	76.0%	94.0%	85.0%	115	115	R	I	1948−1965
Zmijewski (1984)								
Grice and Dugan (2003)	96.3%	36.4%	85.6%	841	183	R	I	1988−1991
Wu et al. (2010)	80.1%	82.0%	80.1%	49724	887	R	I	1980−2006

H: Healthy - F: Failed.
O: results achieved with Original models - R: results achieved with Re-estimated model.
F: Finance - GO: Gas and oil - I: Industry - VS: Various sectors.
Figures presented in this table correspond to the best results when many results were computed.
Empty cells correspond to results that were not mentioned.

38

Table 3: Number of healthy and failed companies by period and by sample

| | Period 1 | | | Period 2 | | | Period 3 | | |
| | ES | TS | | ES | TS | | ES | TS | |
	1998	1999	2000	2000	2001	2002	2002	2003	2004
Healthy	800	900	1100	1100	950	920	920	1000	1050
Bankrupt	800	900	1100	1100	950	920	920	1000	1050
Total	1600	1800	2200	2200	1900	1840	1840	2000	2100
Economic cycle	G	G	G	G	D	D	D	G	G

ES: Estimation sample - TS: Test sample.
G: Growth - D: Downturn.

Table 4: Breakdown of yearly data used for model estimation and test by period

| | Period 1 | | |
	Estimation	Test	
Trajectories, Cox's model, DA 6Y, LR 6Y, NN 6Y	1993 to 1998	1994 to 1999	1995 to 2000
DA 3Y, LR 3Y, NN 3Y	1996 to 1998	1997 to 1999	1998 to 2000
DA 1Y, LR 1Y, NN 1Y	1998	1999	2000
	Period 2		
	Estimation	Test	
Trajectories, Cox's model, DA 6Y, LR 6Y, NN 6Y	1995 to 2000	1996 to 2001	1997 to 2002
DA 3Y, LR 3Y, NN 3Y	1998 to 2000	1999 to 2001	2000 to 2002
DA 1Y, LR 1Y, NN 1Y	2000	2001	2002
	Period 3		
	Estimation	Test	
Trajectories, Cox's model, DA 6Y, LR 6Y, NN 6Y	1997 to 2002	1998 to 2003	1999 to 2004
DA 3Y, LR 3Y, NN 3Y	2000 to 2002	2001 to 2003	2002 to 2004
DA 1Y, LR 1Y, NN 1Y	2002	2003	2004

DA: Discriminant analysis - LR: Logistic regression - NN: Neural network.
1Y: One-year period model - 3Y: Three-year period model - 6Y: Six-year period model.

Table 5: Correct classification rates calculated with data from test samples by period

Period 1 - Estimation sample: 1998 (growth)

| Test samples | DA | | | LR | | | NN | | | Cox | Traj. |
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
1999 (growth)	80.2%	80.7%	79.8%	81.0%	81.1%	79.9%	81.3%	81.2%	78.6%	81.1%	81.2%
2000 (growth)	81.2%	80.9%	78.8%	81.6%	81.1%	79.2%	81.3%	81.0%	76.4%	81.1%	81.9%

Period 2 - Estimation sample: 2000 (growth)

| Test samples | DA | | | LR | | | NN | | | Cox | Traj. |
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
2001 (downturn)	79.2%	79.3%	77.4%	79.9%	79.7%	78.0%	79.9%	79.6%	75.3%	80.4%	80.0%
2002 (downturn)	74.7%	75.1%	73.4%	74.3%	74.8%	72.2%	76.3%	76.6%	72.0%	78.8%	78.7%

Period 3 - Estimation sample: 2002 (downturn)

| Test samples | DA | | | LR | | | NN | | | Cox | Traj. |
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
2003 (growth)	78.3%	77.3%	76.2%	79.9%	79.5%	77.7%	79.7%	79.6%	74.8%	80.0%	80.8%
2004 (growth)	75.5%	75.9%	73.8%	76.0%	76.0%	73.6%	76.9%	76.9%	71.4%	79.5%	80.0%

DA : Discriminant analysis - LR : Logistic regression - NN : Neural network - Cox : Cox's model - Traj.: Trajectories.

Table 6: Significance levels of a test for differences between correct

classification rates achieved with all models and samples from 2000,

2002 and 2004

Models	Samples	Traj.	Cox	NN			LR			DA	
				6Y	3Y	1Y	6Y	3Y	1Y	6Y	3Y
DA 1Y	2000	-	-	*****	-	-	-	-	-	-	-
	2002	*****	*****	*	-	-	-	-	-	-	-
	2004	*****	*****	***	-	-	-	-	-	-	-
DA 3Y	2000	-	-	*****	-	-	-	-	-	-	
	2002	***	***	**	-	-	*	-	-	-	
	2004	*****	***	*****	-	-	-	-	-	-	
DA 6Y	2000	***	-	-	-	*	-	-	**		
	2002	******	******	-	**	*	-	-	-		
	2004	******	******	-	**	**	-	-	-		
LR 1Y	2000	-	-	******	-	-	*	-			
	2002	*****	*****	-	-	-	-	-			
	2004	*****	***	*****	-	-	-	-			
LR 3Y	2000	-	-	**	-	-	-				
	2002	***	*****	-	-	-	-				
	2004	*****	***	-	-	-	-				
LR 6Y	2000	*	-	**	-	-					
	2002	******	******	-	*****	*****					
	2004	******	******	-	**	**					
NN 1Y	2000	-	-	*****	-						
	2002	-	-	*****	-						
	2004	***	**	******	-						
NN 3Y	2000	-	-	*****							
	2002	-	-	*****							
	2004	**	*	******							
NN 6Y	2000	******	*****								
	2002	******	******								
	2004	******	******								
Cox	2000	-									
	2002	-									
	2004	-									

-	p-value > 0.20
*	0.15 < p-value ≤ 0.20
**	0.10 < p-value ≤ 0.15
***	0.05 < p-value ≤ 0.10
****	0.01 < p-value ≤ 0.05
*****	0.001 < p-value ≤ 0.01
******	p-value ≤ 0.001

Table 7: Correct classification rates calculated with data from test

samples by period and company status (healthy vs. failed)

Period 1 - Estimation sample: 1998 (growth)					
Test samples	DA	LR	NN	Cox	Traj.

		DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
1999 (growth)	Healthy	81.1%	81.8%	81.3%	81.7%	81.3%	82.0%	82.2%	81.3%	81.7%	81.0%	80.4%
	Failed	79.3%	79.6%	78.2%	80.3%	80.8%	77.8%	80.3%	81.1%	75.6%	81.2%	81.9%
2000 (growth)	Healthy	84.5%	83.6%	84.7%	83.4%	83.2%	83.5%	82.6%	81.8%	83.7%	81.9%	80.3%
	Failed	77.9%	78.1%	72.9%	79.9%	79.1%	74.9%	79.9%	80.1%	69.0%	80.3%	83.5%

Period 2 - Estimation sample: 2000 (growth)

Test samples		DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
2001 (downturn)	Healthy	82.1%	82.7%	83.7%	81.9%	82.4%	81.6%	83.7%	81.6%	81.4%	82.3%	79.3%
	Failed	76.2%	75.9%	71.1%	77.9%	76.9%	74.4%	76.1%	77.7%	69.2%	78.5%	80.7%
2002 (downturn)	Healthy	86.0%	85.8%	86.7%	84.5%	80.1%	80.5%	86.0%	79.6%	86.0%	84.0%	77.3%
	Failed	63.5%	64.3%	60.0%	64.1%	69.6%	63.8%	66.6%	73.6%	58.0%	73.5%	80.1%

Period 3 - Estimation sample: 2002 (downturn)

Test samples		DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
2003 (growth)	Healthy	80.1%	79.5%	82.7%	81.7%	81.2%	83.5%	82.1%	82.7%	84.6%	80.3%	79.8%
	Failed	76.4%	75.1%	69.6%	78.0%	77.7%	71.9%	77.3%	76.5%	64.9%	79.6%	81.7%
2004 (growth)	Healthy	83.9%	82.5%	84.4%	83.5%	81.7%	83.0%	83.9%	79.6%	82.4%	81.2%	79.8%
	Failed	67.0%	69.3%	63.1%	68.6%	70.2%	64.2%	69.8%	74.1%	60.5%	77.7%	80.1%

Table 8: Significance levels of a test for differences between correct

classification rates achieved with all models, samples from 2000,

2002 and 2004, and calculated with solely healthy firms

Healthy firms

Models	Samples	Traj.	Cox	NN			LR			DA	
				6Y	3Y	1Y	6Y	3Y	1Y	6Y	3Y
DA 1Y	2000	*****	**	-	**	-	-	-	-	-	-
	2002	******	-	-	******	-	******	******	-	-	-
	2004	*****	**	*****	*****	-	-	-	-	-	-
DA 3Y	2000	***	-	-	-	-	-	-	-	-	
	2002	******	-	-	******	-	******	******	-	-	
	2004	**	-	**	**	-	-	-	-	-	
DA 6Y	2000	*****	**	-	***	-	-	-	-		
	2002	******	**	-	******	-	******	******	*		
	2004	*****	***	******	******	-	-	**	-		
LR 1Y	2000	***	-	-	-	-	-	-			
	2002	******	-	-	******	-	*****	*****			
	2004	*****	*	*****	*****	-	-	-			
LR 3Y	2000	**	-	-	-	-	-				
	2002	**	*****	******	-	******	-				
	2004	-	-	***	-	-	-				
LR 6Y	2000	***	-	-	-	-					
	2002	***	***	******	-	******					
	2004	***	-	***	***	-					
NN 1Y	2000	*	-	-	-						
	2002	******	-	-	******						
	2004	*****	**	*****	*****						
NN 3Y	2000	-	-	-							
	2002	-	*****	******							
	2004	-	-	-							
NN 6Y	2000	***	-								
	2002	******	-								
	2004	-	-								
Cox	2000	-									

Table 9: Significance levels of a test for differences between correct classification rates achieved with all models, samples from 2000, 2002 and 2004, and calculated with solely failed firms

Failed firms

Models	Samples	Traj.	Cox	NN			LR			DA	
				6Y	3Y	1Y	6Y	3Y	1Y	6Y	3Y
DA 1Y	2000	******	*	******	-	-	**	-	-	*****	-
	2002	******	******	*****	******	*	-	******	-	**	-
	2004	******	******	******	******	*	-	*	-	***	-
DA 3Y	2000	******	-	******	-	-	**	-	-	*****	
	2002	******	******	******	******	-	-	*****	-	***	
	2004	******	******	******	*****	-	*****	-	-	******	
DA 6Y	2000	******	******	***	******	******	-	******	******		
	2002	******	******	-	******	******	***	******	***		
	2004	******	******	-	******	******	-	******	*****		
LR 1Y	2000	*****	-	******	-	-	*****	-			
	2002	******	******	******	******	-	-	*****			
	2004	******	******	******	******	-	*****	-			
LR 3Y	2000	*****	-	******	-	-	*****				
	2002	******	***	*****	***	*	******				
	2004	******	******	**	***	-	******				
LR 6Y	2000	******	******	******	******	*****					
	2002	******	******	*****	******	-					
	2004	******	******	**	******	*****					
NN 1Y	2000	*****	-	******	-						
	2002	******	******	******	******						
	2004	******	******	******	*****						
NN 3Y	2000	***	-	******							
	2002	******	-	******							
	2004	******	***	******							
NN 6Y	2000	******	******								
	2002	******	******								
	2004	******	******								
Cox	2000	*****									
	2002	******									
	2004	*									

Table 10: Percentage of misclassified firms, calculated with test
sample from 2000, by model for different misclassification costs

Test sample: 2000

Costs (c_1)	Error types	Models DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
	Type-I	98.0%	99.5%	100.0%	89.9%	90.6%	100.0%	89.6%	95.6%	100.0%	81.0%	76.3%
1	Type-II	1.6%	0.0%	0.0%	1.1%	1.6%	0.0%	1.9%	1.2%	5.9%	7.1%	1.7%
	Total	49.8%	49.7%	50.0%	45.5%	46.1%	50.0%	45.8%	48.4%	53.0%	44.0%	39.0%
	Type-I	78.8%	80.1%	76.1%	64.7%	65.0%	87.7%	64.4%	60.5%	78.8%	61.4%	57.1%
10	Type-II	5.2%	6.9%	6.1%	8.1%	8.4%	7.2%	7.4%	7.0%	6.9%	9.9%	10.7%
	Total	42.0%	43.5%	41.1%	36.4%	36.7%	47.5%	35.9%	33.7%	42.9%	35.6%	33.9%
	Type-I	42.5%	44.3%	44.2%	34.1%	41.4%	44.8%	34.3%	32.3%	45.4%	29.0%	26.8%
30	Type-II	12.0%	13.4%	12.5%	13.6%	12.2%	13.6%	14.0%	16.8%	10.4%	15.0%	16.2%
	Total	27.2%	28.8%	28.4%	23.9%	26.8%	29.2%	24.1%	24.5%	27.9%	22.0%	21.5%
	Type-I	22.1%	21.4%	26.8%	19.5%	20.7%	24.9%	19.9%	18.1%	30.6%	19.4%	16.3%
50	Type-II	15.9%	16.6%	15.3%	16.8%	17.0%	16.8%	17.8%	18.3%	16.7%	18.5%	20.0%
	Total	19.0%	19.0%	21.0%	18.2%	18.9%	20.9%	18.9%	18.2%	23.7%	18.9%	18.1%
	Type-I	17.0%	16.1%	20.7%	10.1%	10.8%	20.5%	9.7%	8.2%	19.4%	7.4%	6.1%
70	Type-II	26.9%	21.9%	24.4%	22.5%	23.6%	25.2%	22.6%	23.7%	27.7%	22.2%	20.8%
	Total	22.0%	19.0%	22.5%	16.3%	17.2%	22.8%	16.2%	16.0%	23.5%	14.8%	13.5%

Table 11: Percentage of misclassified firms, calculated with test
sample from 2002, by model for different misclassification costs

Test sample: 2002

Costs (c_1)	Error types	Models DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
	Type-I	94.0%	98.7%	100.0%	96.8%	97.8%	100.0%	96.7%	95.7%	100.0%	83.0%	67.5%
1	Type-II	8.8%	10.3%	6.8%	4.1%	4.9%	1.8%	8.3%	2.5%	8.3%	3.7%	2.3%
	Total	43.0%	45.6%	44.7%	42.2%	43.0%	42.6%	43.9%	41.0%	45.3%	36.3%	29.2%
	Type-I	76.1%	86.7%	84.2%	76.3%	85.3%	91.8%	84.3%	79.6%	89.7%	77.4%	64.1%
10	Type-II	9.7%	11.1%	9.9%	12.6%	14.2%	8.5%	9.9%	10.9%	9.9%	8.5%	9.8%
	Total	35.9%	40.9%	39.4%	37.2%	41.6%	42.0%	39.4%	37.8%	41.6%	35.9%	30.9%
	Type-I	55.5%	58.9%	59.2%	59.3%	55.0%	60.3%	52.7%	48.4%	61.6%	47.2%	38.7%
30	Type-II	10.7%	12.8%	11.3%	13.9%	18.0%	16.7%	12.1%	13.8%	12.8%	13.9%	15.1%
	Total	27.7%	30.0%	29.5%	30.6%	30.5%	32.2%	27.1%	26.0%	31.1%	25.5%	22.5%
	Type-I	36.0%	35.3%	39.7%	35.3%	29.5%	35.5%	33.2%	26.1%	41.1%	25.9%	19.6%
50	Type-II	14.3%	14.9%	13.6%	15.7%	20.7%	19.9%	14.0%	20.8%	14.6%	16.1%	22.8%
	Total	21.0%	21.0%	22.3%	21.3%	21.0%	23.2%	19.7%	19.6%	23.3%	17.5%	17.7%
	Type-I	27.6%	26.6%	30.0%	28.7%	25.2%	29.5%	24.6%	17.0%	30.2%	21.8%	12.7%
70	Type-II	31.1%	30.2%	33.9%	31.2%	31.7%	35.0%	26.7%	28.8%	35.0%	27.0%	29.2%
	Total	24.5%	23.8%	26.7%	25.0%	23.8%	27.0%	21.5%	19.1%	27.3%	20.4%	17.5%

Table 12: Percentage of misclassified firms, calculated with test
sample from 2004, by model for different misclassification costs

Test sample: 2004

Costs	Error	Models

43

(c_1)	types	DA			LR			NN			Cox	Traj.
		1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
	Type-I	98.8%	95.2%	100.0%	97.4%	95.3%	99.9%	94.1%	93.9%	100.0%	80.7%	66.8%
1	Type-II	4.9%	7.4%	10.0%	4.2%	4.0%	11.1%	3.5%	3.2%	7.4%	2.9%	3.0%
	Total	49.5%	49.0%	52.5%	48.5%	47.4%	53.0%	46.6%	46.4%	51.3%	39.9%	33.3%
	Type-I	89.0%	85.7%	92.9%	86.5%	83.3%	87.9%	76.5%	75.1%	93.4%	66.8%	59.3%
10	Type-II	5.5%	8.3%	10.7%	6.2%	5.8%	11.9%	5.3%	5.0%	8.1%	6.1%	6.9%
	Total	45.1%	44.9%	49.4%	44.2%	42.5%	47.6%	39.0%	38.3%	48.5%	34.8%	31.6%
	Type-I	49.8%	49.0%	54.9%	48.7%	46.5%	52.2%	44.3%	42.5%	57.4%	38.4%	34.8%
30	Type-II	9.7%	10.2%	14.7%	9.4%	9.7%	15.9%	9.3%	8.2%	13.1%	9.6%	11.7%
	Total	28.4%	28.2%	33.2%	27.7%	26.8%	32.5%	25.6%	24.2%	33.7%	22.9%	22.2%
	Type-I	32.6%	30.3%	36.4%	31.0%	29.2%	34.8%	29.4%	25.2%	39.0%	21.7%	19.9%
50	Type-II	16.2%	18.2%	16.7%	16.8%	18.6%	17.2%	16.5%	20.4%	17.8%	19.1%	20.5%
	Total	23.3%	23.1%	25.3%	22.8%	22.8%	24.8%	21.9%	21.8%	27.1%	19.5%	19.3%
	Type-I	24.2%	22.8%	24.5%	21.8%	22.0%	23.1%	21.5%	20.7%	27.3%	17.0%	13.8%
70	Type-II	28.2%	31.0%	26.6%	24.9%	28.8%	27.9%	27.3%	30.1%	28.1%	25.8%	27.4%
	Total	25.0%	25.6%	24.4%	22.3%	24.2%	24.4%	23.3%	24.2%	26.5%	20.4%	19.7%

Table 13: Resubstitution risks by model, calculated with test sample from 2000, for different misclassification costs

Test sample: 2000

Costs (c_1)	Models									Cox	Traj.
	DA			LR			NN				
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
1	0.036	0.020	0.020	0.029	0.034	0.020	0.037	0.031	0.078	0.086	0.032
10	0.208	0.228	0.212	0.209	0.212	0.246	0.201	0.190	0.225	0.220	0.219
20	0.320	0.353	0.330	0.296	0.307	0.355	0.296	0.290	0.305	0.277	0.273
30	0.372	0.397	0.388	0.338	0.368	0.403	0.343	0.358	0.374	0.321	0.319
40	0.407	0.417	0.404	0.347	0.352	0.429	0.352	0.364	0.427	0.341	0.335
50	0.377	0.377	0.418	0.360	0.374	0.414	0.374	0.360	0.470	0.374	0.359
60	0.456	0.404	0.473	0.491	0.487	0.476	0.378	0.348	0.586	0.341	0.327
70	0.502	0.440	0.529	0.361	0.383	0.533	0.358	0.347	0.543	0.320	0.289

Resubstitution risk $= p_1 c_1 e_1 / N_1 + p_2 c_2 e_2 / N_2$,
where p_1 and p_2 are the respective prior probabilities of bankruptcy and non-bankruptcy; e_1 and e_2 are the respective type-I (failed firm predicted as healthy) and type-II (healthy firm predicted as failed) errors; N_1 and N_2 are the respective numbers of failed and healthy firms in the sample and c_1 and c_2 are the respective costs of type-I and type-II errors.

Table 14: Resubstitution risks by model, calculated with test sample from 2002, for different misclassification costs

Test sample: 2002

Costs (c_1)	Models									Cox	Traj.
	DA			LR			NN				
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
1	0.105	0.121	0.087	0.060	0.068	0.038	0.100	0.044	0.101	0.053	0.036
10	0.247	0.282	0.265	0.276	0.310	0.267	0.266	0.266	0.276	0.238	0.224
20	0.371	0.408	0.409	0.405	0.427	0.447	0.376	0.369	0.405	0.371	0.345
30	0.438	0.479	0.466	0.492	0.507	0.526	0.435	0.426	0.495	0.419	0.380
40	0.472	0.496	0.495	0.475	0.513	0.531	0.464	0.412	0.532	0.423	0.357
50	0.500	0.499	0.530	0.507	0.497	0.550	0.469	0.464	0.554	0.416	0.419

	DA			LR			NN			Cox	Traj.
60	0.598	0.605	0.635	0.619	0.585	0.670	0.559	0.490	0.568	0.496	0.440
70	0.691	0.669	0.752	0.707	0.664	0.755	0.606	0.520	0.766	0.570	0.465

Table 15: Resubstitution risks by model, calculated with test sample from 2004, for different misclassification costs

Test sample: 2004											
Costs						Models					
(c_1)		DA			LR			NN		Cox	Traj.
	1Y	3Y	6Y	1Y	3Y	6Y	1Y	3Y	6Y		
1	0.067	0.092	0.118	0.061	0.058	0.129	0.053	0.051	0.093	0.044	0.042
10	0.232	0.253	0.290	0.234	0.224	0.292	0.205	0.200	0.266	0.193	0.186
20	0.363	0.385	0.441	0.368	0.357	0.429	0.321	0.294	0.442	0.302	0.304
30	0.394	0.394	0.473	0.384	0.374	0.469	0.357	0.335	0.473	0.325	0.323
40	0.439	0.463	0.470	0.425	0.439	0.487	0.380	0.387	0.497	0.355	0.346
50	0.484	0.481	0.527	0.474	0.474	0.517	0.456	0.452	0.565	0.405	0.400
60	0.568	0.577	0.610	0.544	0.527	0.608	0.509	0.515	0.622	0.447	0.439
70	0.615	0.622	0.603	0.549	0.590	0.597	0.569	0.584	0.658	0.490	0.462