A QUANTITATIVE ANALYSIS OF TRIGGER FACTORS INFLUENCING COMPANIES FAILURE IN ITALY

Francesca di Donato, Assistant Professor of Accounting, PhD
Luciano Nieddu, Assistant Professor of Statistics, PhD
UNINT – Universita degli Studi Internazionali di Roma, Italy

Abstract

   In this paper the problem of firms' failures will be considered. The aim is to determine which are the trigger factors that can predict the inability of a firm to cover its obligations. The problem will be tackled both from a cross-sectional and a longitudinal point of view using non parametric and semi-parametric statistical models that allow to gain information from the data without making too many assumptions.

Keywords: Cox Model, survival analysis, outliers, business failure

Introduction

   An important body of research in accounting has focused on the use of financial ratios to predict firms' bankruptcy. The use of financial ratios and the development of empirical approach to predict companies' failure have a long history. Since Beaver's (1966) and Altman's (1968) pioneering works, many studies have been devoted to exploring the use of accounting information in predicting business failure.

   Bankruptcy occurs when firms lack sufficient capital to cover the obligation of a business (Boardman et al. 1981). Beaver (1966) defines failure as the inability of a firm to pay its financial obligations as they mature and establishes that financial ratios have predictive power up to at least five years prior to bankruptcy. Therefore, the initial approaches have been focused on using financial ratios at a given time prior to the occurrence of the event to determine the probability of bankruptcy. No effort has been made to use longitudinal information.

   Beaver (1966) uses only univariate statistics on US market data to determine the effect of financial ratios on the probability of bankruptcy. As noted by Altman (1968), although the univariate approach establishes certain important generalizations regarding the performance and trends of particular measurements, the adaption of their results for assessing bankruptcy potential

of firms, both theoretically and practically, is questionable. The univariate nature focuses on individual signals of impending problems.

Altman in 1968 introduced the use of Multiple Discriminant Analysis (MDA) to predict whether a firm will experience bankruptcy. While multivariate in nature, this analysis does not take into account the evolution of the financial ratio over time. In 1980 Ohlson, (Ohlson, 1980) suggested the use of conditional logistic regression to overcome some of the shortcomings of MDA, using information of the performance of each firm at various stages prior to bankruptcy.

Only in the mid 1980s, the studies started to focus on the use of longitudinal models and semi-parametric and non parametric approaches such as recursive partitioning algorithms (Frydman et al., 1985, Hastie et al., 2001), neural networks techniques (Odom and Sharda, 1990; Coats and Fant, 1992; Tam and Kiang, 1992; Wilson and Sharda, 1994) and survival analysis (Cox, 1972, Lane, Looney and Wansley, 1986; Crapp and Stevenson, 1987; Chen and Lee, 1993; Bandopadhyaya, 1994; Therneau and Grambsch, 2000).

Despite the increasing use of survival analysis (SA) in modeling financial distress, little attention has been given to the use of time varying covariates to estimate these models, that started to attract increasing attention following the dynamic model of Shumway (2001). Shumway considered longitudinal data and a semi-parametric model to determine the probability of failure of a firm, allowing for time dependent covariates to influence the hazard function, defined as the probability of a firm to experience bankruptcy at time t given the fact that it has survived until that time. The use of time dependent covariate allows the varying financial indicator to vary their effect on the probability of bankruptcy, therefore yielding a dynamic model. More recent studies using the hazard function are Romer (2005), Chanchrat et al. (2007), Kim and Parkington (2008), Nieddu and Vitiello (2013).

The ability of SA models to account for time varying effects appear to be more suited to modeling a dynamic process, such as business failure, than cross-sectional models. This could results in a better predictive accuracy of SA models when compared with cross-sectional models such as linear and nonlinear discriminant analysis. Another interesting feature of SA when compared with discriminant analysis, is that it doesn't assume that the data come from two different populations but rather assumes that all businesses come from the same population distribution. In SA models, the successful businesses are distinguished by treating them as censored data, which indicates that their time of failure is not yet known. This assumption more accurately models the real world (Laitinen and Luoma, 1991).

Furthermore, SA does not make any of the restrictive distribution assumptions inherent in DA and LA, such as linearity. The semi-parametric

and parametric SA models make some distribution assumptions, but they are not so commonly violated.

In this paper we use two different techniques to investigate the phenomenon of firm failures. Namely we study the phenomenon from a cross-sectional point of view using a non parametric approach (classification trees) to determine the conditional probability of bankruptcy of a firm at various time prior to occurrence of the event. Together with this static analysis we use a longitudinal analysis to determine the influence of varying covariates on the hazard function via a proportional Cox model. One of the strength of the Cox model for SA is its ability to take into account covariates that change over time.

The two approaches will be applied on original data collected over a decade (2000-2010) for a stratified sample of non listed Italian companies. The reasons for a double analysis are twofold: firstly, we want to determine which are the financial statement components that influence bankruptcy at various point in time using a robust non parametric technique which allows to mine the information on the data without requiring any prior assumption.

Secondly, after verifying the existence of a relationship between the covariates and the failure occurrence, we want to analyze the influence over time of financial ratios on the hazard function, taking into account the effect of variations of these indicators on the risk function. In order to do that we will use a very flexible semi parametric model, such as the Cox proportional hazard model. It must be stressed out that for the meaningfulness of the results, the proportional hazard hypothesis must be verified.

This paper differs from analogous papers on the topic for the following reasons. First of all, we use a non parametric technique to test if there is a real relation between data at hand and firms survival. Once we have established such a link, we first test the possibility of applying semi parametric approach such as the Cox model and this step has been neglected in all the literature we analyzed. This is not a trivial point since the results of the Cox model are only meaningful if the proportional hazard hypothesis holds. Lack of robustness of the Cox model from departure from proportionality and in the presence of influential outliers has been stressed in the literature (see Bednarsky, 1989, Cain and Lang 1984). We have handled anomalous observations taking into account their effect without dropping them out of the study.

Moreover we have used a stratified random sample using business sectors as stratifying variable selecting only firms with revenues from sales from euro 2 millions to 50 millions.

The results concern a retrospective study since the aim of the paper is not to determine the proportion of failed firms but to determine the factors affecting the failure. Therefore 50 active firms and 50 failed firms have been

selected and their financial statements have been studied during a period of 10 years.

The layout of the paper is as follows. In Section 2 the various methodologies will be described. Section 3 presents the empirical results concerning the application of the classification trees while Section 4 will present the results of the Cox model. Finally in Section 5 some conclusion will be drawn.

Statistical Models:

A. Survival Analysis

The problem of analyzing time to event data arises in a number of applied fields, such as medicine, biology, public health, epidemiology, engineering, economics, and demography. Four functions characterize the distribution of the time until some specified event occurs, namely, the survival function, which is the probability of an individual surviving to time t; the hazard rate function, which is the chance an individual living at time t to experiences the event in the next instant in time; the probability density function, which is the unconditional probability of the event's occurring at time t. The interpretations of the survival function and the hazard function is very different, but either one can be derived from the other. Although the hazard function must be non negative and its integral over $\mathbb{R}^+$ must be plus infinite, other than these it has no other constraints.

There are many different SA techniques to estimate the survival and hazard functions. The most popular of these is a non-parametric technique known as the Product-Limit, or Kaplan-Meier, estimator (Kaplan Meier, 1958). The Kaplan-Meier estimate of the survival function $S(t)$ corresponds to the non-parametric maximum likelihood estimate of $S(t)$.

Many parametric models that can be used to asses the importance of various covariates in the survival time through the hazard function are available.

In survival analysis dependence of survival time from covariates is expressed modeling the hazard function $\lambda(t, x)$:

$$\lambda(t, x) = \lim_{\Delta t \to 0} \frac{\Pr(t < T < t + \Delta t | T > t)}{\Delta t}$$

Due to its flexibility, the Cox's Proportional Hazard model (Cox, 1972) is the most applied in the medical and business failure field. It is a semi-parametric model and is defined as:

$$\lambda(t, x) = \lambda_0(t) e^{x\beta}$$

where $\lambda_0(t)$ is an arbitray unspecified function of time, is termed the baseline hazards function; it is the hazard for an individual with all the covariates set to zero and describes how the hazard function changes over time. It is the non-parametric part of the model. The linear predictor $x\beta$ is a time

independent quantity that describes how the hazard function relates to the business specific explanatory variables and is the parametric part of the model. Note that some or all of the explanatory variables can be time dependent.

The likelihood can be written as:

$$L(\beta) = \prod_i \lambda(t, x)^{w_i} S(t_i)$$

where w_i is a censoring indicator that takes value 1 if the survival time t for the i-th observation is uncensored and zero if it is censored. Cox used a conditioning argument to eliminate the baseline hazard in a partial likelihood framework. This partial likelihood could then be maximized without reference to the unknown baseline hazard

$$L_p(\beta) = \prod_i \left\{ \frac{e^{x_i \beta}}{\sum_{j \in R_i} e^{x_j \beta}} \right\}^{w_i}$$

where R_i is the set of individuals at risk at time t_i. The resulting maximum partial likelihood estimator is consistent and asymptotically normal (Andersen and Gill, 1982).

The Cox model assumes that time is continuous so ties events cannot occur and if they happen it is because of the time-measuring device. There are several ways to deal with ties, see Breslow (1974) and Efron (1977) for instance. In the following the Efron's approach to handle ties will be used.

The validity of Cox's regression analysis relies heavily on the assumption of proportionality of the hazard rates of individuals with distinct values of a covariate. If the proportional hazard assumption is violated for a variable, then, one approach to dealing with this problem is to stratify on this variable. Stratification fits a different baseline hazard function for each stratum, so that the form of the hazard function for different levels of this variable is not constrained by their hazards being proportional. The effect of the non-stratifying covariates is assumed to be the same across strata, although it is possible to extend the model to consider interaction between covariate and strata. It is assumed, however, that the proportional hazards model is appropriate within strata for the other covariates. This approach is not free of drawbacks: for instance the baseline hazards are estimated separately for each stratum, i.e. on a reduced sample size; we lose the possibility to quantify the effect of the stratifying variable and it is not possible for continuous covariates unless a cut point is arbitrarily selected.

Nonetheless, regardless of some shortcomings, the built in time factor in SA models allows them to model time-dependent explanatory variables. Zavgren (1985) found that in business failures the signs of the explanatory variable coefficients may change considering the indicators at various stages in time prior to the occurrence of the event. Thus, an advantage of SA is the

capability to model these changes, which cannot be done with cross-sectional models.

B. *Classification trees*

Classification trees are a non-parametric supervised learning method used in data mining for patter recognition and classification. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data at hand that has been previously classified by an expert or by any other way deemed appropriate.

Tree-based methods partition the covariate space X into disjoint set of rectangular regions, and then classify the observations according to which partition element they fall in. The partitioning is usually performed according to an impurity measure (usually the Gini index) or according to the information gain (entropy) that can be achieved once the covariate space has been partitioned. Therefore, starting with a single node (root) we look for the binary partition that yields the most information about the class.

This partitioning is recursively performed on the derived subsets and it stops either when the units in a node have all the same value of the variable indicating the class or when splitting no longer adds value to the predictions. The iterative partitioning process is called "growing a tree" or "learning".

When there are several covariates, we choose whichever covariate and split that leads to the lowest impurity. This process is continued until some stopping criterion is met. For example, we might stop when every partition element has less than a certain number of elements. The bottom nodes of the tree are called the leaves. Each leaf is assigned a class according to a majority rule based on the classes of the elements that belong to that leaf. This majority rule criterion is also used in classifying new objects.

Decision-tree learners can create over-complex trees. The complexity of the tree doesn't necessarily imply a good accuracy of the tree. To avoid overcomplex trees, pruning techniques usually based on cross validation (i.e. on their performance on new data) can be used.

Classification trees have only been applied once to business failure in a study that did not produce reliable results due to a very small sample size (Huarng et al., 2005).

Results

CLASSIFICATION TREES

In this paragraph the results of the application of classification trees on the data available will be presented. All the results have been obtained using the package rpart() (recursive partitioning and regression trees) that follows the approach from Breimans et al (1984). The package is available for the statistical software R (www.r-project.org) that has been used for all the analysis.

The first part of the study is a cross sectional analysis where we have considered firms at various years prior to failure. Each firm that experienced a failure has been matched with an analogous one from the same year that is still active at the end of the study. The analysis has been carried out for up to 8 years prior to failure because going any further would have decreased the sample size too much. Only the most significant results will be displayed in this paragraph.

All the main financial statement items and the performance indicators have been used as explanatory variabile to assess the probability of failure. The financial items for each firm have been normalized dividing them by the average sales of firms with the same size of the one considered, to make all the items comparable. A pruning of each tree has been carried out to avoid overifitting.

In Table I the various type of errors for the pruned trees for each year have been displayed. The resubstitution error is a biased estimate of the performance of the tree and can be used to check how well the tree fits the data. The xerror and che cross validation error (cv error in the table) are correct estimates of the performance of the tree and are obtaiened using cross validation, therefore they are more suited to evaluate the ability of tree to classify new firms. It is evident from the Table that the further we go back in time, the harder it becomes for the tree to assess the outcome of the firm.

PERFORMANCE PARAMETER OF THE FITTED TREES AFTER PRUNING

years prior to failure	resubstitution error	xerror	cv error
1	0.2600	0.6000	0.3000
2	0.5000	0.6000	0.3000
3	0.4200	0.7600	0.3800
4	0.4286	0.4286	0.2143
5	0.4565	0.8044	0.4022
6	0.3030	0.6061	0.3030
7	0.2609	1.0000	0.5000
8	0.4118	0.7647	0.3824

In Figure 1 the classificationt tree for firms 5 years prior to failure has been displayed. The main item in the financial statement relevant to discriminate between failure and survival 5 years prior to the event is equity. Namely very high values of relative equity tend to garantie a survival (31 active over 10 failed firms). A low value of equity, on the other hand, is more common for failed companies (15 active over 36 failed). To further refine the classification, financial debts on working capital and FI/Ebida could be used for firms with high equity and low equity respectively.

In Figure 2 the classifcation tree for firms 1 year prior to the event has been displayed. Liquidity is now the discriminant variabile to classify a firm as active or failed. Very high value of liquidity tend to be associated with active firms (41 active over 18 failed firms) while very low values tend to be associated with failed frims (9 active over 32 failed). The classification can be further refined considering financial interests on sales, sales and longterm liabilities.

In all the models that have been fitted, the variables that have been constantly influencing the risk of failure, providing the higher information gain as a first split in the classification tree, have been equity and liquidity which are structural variables representing the solidity of the company.

Both variables are very important for a company because equity, which is shareholders capital, is the best financing source as it does not generate interest expenses and ensures high strength and low risk of failure.

In particular, equity is very important in terms of firms failure in the years long standing from bankruptcy, mainly 8, 6, 5 and 3 years before bankruptcy. This highlights the importance of this kind of source for Italian small and medium-sized enterprises, in order not to be dependent on the market and in particular on the banks.

Instead, 4 and 2 years prior to bankruptcy, the discriminant variable for a firm is represented by the amount of liquidity, which highlights the importance of having the necessary resources to meet payments. A company with high liquidity is autonomous in paying financial and operating obligations and does not need to resort to other costly financial funds.

The existence of a high degree of liquidity is moreover a sign of companies' health because firms with a lot of cash regularly collect trade receivables and therefore do not have problems of uncollectable. But a high degree of liquidity can also depend on a good access to credit, meaning that banks trust the companies and lend them money, meaning that they are healty and solvent

The importance of liquidity variable is consistent with the global financial crisis that has produced many problems for small-medium companies that are not very capitalized. They suffered from a situation of illiquidity due to the difficulties to collect receivables from customers and due to the credit crunch by the banks drastically reducing their fundings due to the lack of trust..
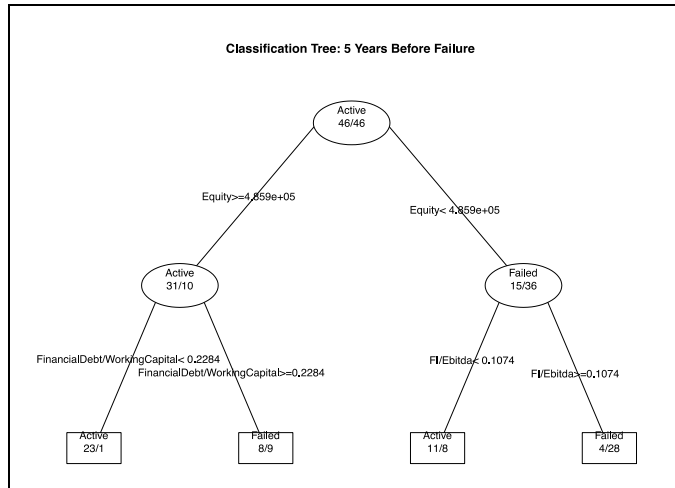
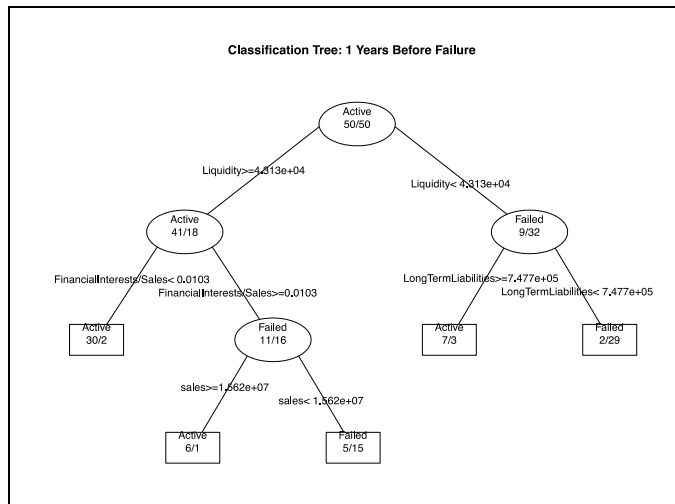Figure 1: classification tree for 5 years prior to failure

Figure 2: classification tree for 1 year prior to failure

SURVIVAL ANALYSIS: RESULTS

In this section the results of the SA will be presented. All the results have been obtained using the survival package for R. The performed analysis will consist in two parts:

- a time-to-event study of the bankruptcy over the period 2000-2010
- an analysis of the age of the firm at the moment of bankruptcy

They are both retrospective studies and they vary for the definition of the outcome. In Figure 3 the Kaplan Meier estimates of the survival function for the year of failure controlling for firm size have been displayed. There is an increase in the rate of failure expecially for micro companies in the years 2005 and 2008. The first drop can be explained by the increase in efficiency

of the iter required to register a bankruptcy due to a change in Italian legislation about failures in 2004. The second drop is due to the international financial and ecomic crisis affecting all companies with higher disavantages for micro firms. In Figure 4 the Kaplan-Meier estiamates for age of the firm at failure, controlling for size of the firm, have been displayed. Overall the two figures highlight a different behaviour of the time to event depending on different size of companies, therefore, in the future models, size will be considered as a confounding factor and it will be controlled for in the analysis.
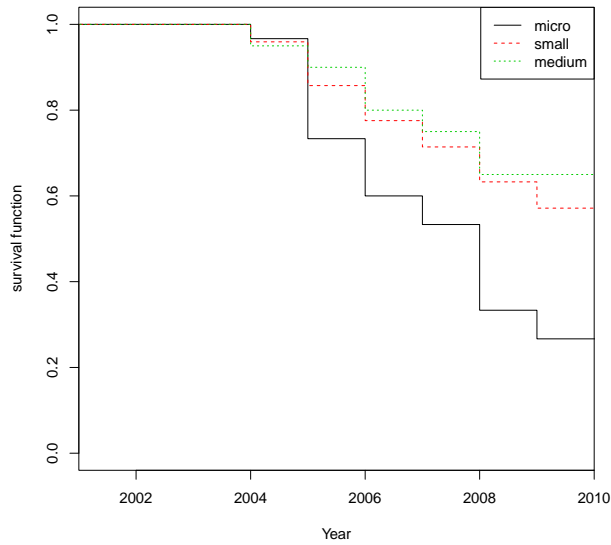

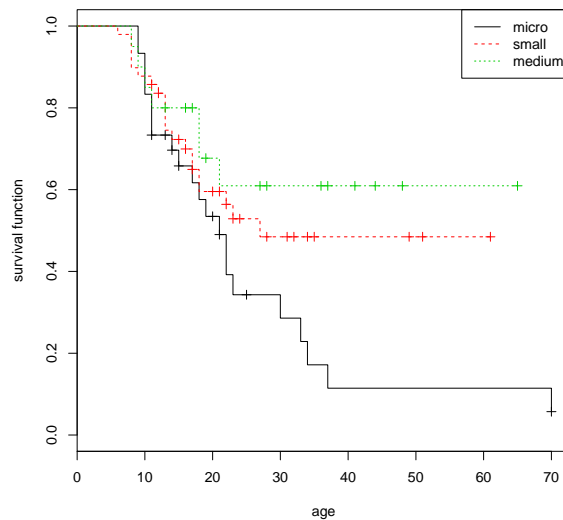Figure 3: Kaplan-Meier estimate for the year of bankruptcy by size of the firm

Figure 4: Kaplan-Meier estimate for the age at bankruptcy by size of the firm

10

From the results from the previous section and from the Kaplan-Meier estimates it is clear that there is a difference in behavious regarding the risk of failure at various points in time for firms. These differences can be partially explained by the size of the firm and by the financial performance of the firm. In the following we will study how the effect on the hazard of the various performance indicators/ratios can be quantified. In order to do so, a Cox model has been applied to the age and time of failure respectively.

A key point that need to be stressed is the causality issue: we would like to study the effect of the performance indicator in predicting the time to event. Consideing the values of the indicator at the year of failure could cause some problems since it could be argued that they are the effect of the failure and not the cause. It would then be safer to consider lagged covariates using the values of the indicators at time t-1 (in years).

The covariates that will be used to model the hazard can be classified into two groups: economics and financial covariates. The list of indicators that will be used to model the hazard function has been displayed in Table II together with some basic descriptive statistics computed over the whole period of time considered in the analysis (2000-2010).

DESCRIPTIVE STATISTICS (25TH PERCENTILE, MEAN, MEDIAN, 75TH PERCENTILE)

Indicators	Descriptive Statistics			
	Q1	Mean	Median	Q3
Return on equity (ROE)	0,00	0,12	0,06	0,28
Return on investment (ROI)	0,00	-0,40	0,07	0,19
Return on sales (ROS)	0,01	-0,29	0,03	0,06
Capital Turnover (CT)	0,42	-6,62	2,10	5,12
Ebitda on Sales (ES)	0,02	-0,23	0,06	0,11
Financial debts on equity (FDE)	0,00	5,32	0,47	3,81
Financial Interest on Ebitda (FIE)	0,01	0,20	0,17	0,43
StructureRatio1 (ST1)	0,15	17,21	0,51	1,26
StructureRatio2 (ST2)	0,61	39,34	1,21	2,71
Working Capital Cycle (WCC)	-29,62	17,22	8,76	64,92
Financial Debts on Working Capital (FDWC)	0,00	0,35	0,10	0,37

Q1=25th percentile; Q3=75th percentile

Since, from the initial analysis of the available data some of the data observed values seem to be affected by anomalies and clearly show outliers, we have decided the recode each ratio considering into three variables: two dummy variables indicating if the ratio is either too high or too low and a variable indicating the value actually assumed by the ratio in case it assumes a value in the normal range. This has been necessary to avoid the distortion

that can be caused by the presence of outliers in the estimates of the Cox model without having to reduce the sample size.

For instance, for return on investment (ROI) the normal range has been set to [0; 0.5]; values lower than 0 will be considered low (dummy variable ROIA=1) and values greater than 0.5 will be considered high (dummy variabile ROIB=1).

For each model the preliminary proportionality test has been performed based on the scaled Schoenfeld residuals. In all the models, the test doesn't allow to refuse the proportional hazard hyphotesis, except in the model using ROE as covariate. When the non proportionality assumption was not verified a stratified model has been fit to the data stratifing with respect to the variabile violating the assumption.

In Table III the exponential of the estimates of the parameter in the Cox model (hazard ratios) for year as time to event have been displayed. For each performance indicator the estimates of $\exp(\beta)$ have been depicted. When the estimates were not signifcantly different from zero, "ns" has been displayed, and when it was significantly different, the estimate together with the level of significance have been reported. Almost all the economic ratios are not significant, except for ROS and ES, but ROS doesn't seem to influence the relative hazard when it assumes normal values or very high values, while it more than doubles (2.07) the risk of failure when it is lower than zero (ROSA=1).

With respect to ES, an increase of one unit of ES reduces the relative hazard by 99%. A value of ES greater than 1 (ESB=1) increases the relative risk by 7. It looks counterintuitive but in this case a too high value of ES must be read as an outlier and so a sympthom of something that is wrong in the financial statement.

Concerning the financial ratios, ST1 and ST2 are never significant except for anomalous values which, respectively, increase the risk of failure by nearly three times for very low values of ST1 (2.92) and reduce the risk of failure by nearly 70% for very high values of ST2 (0.28).

For FDE and FIE, very high values have no significant effect on the relative risk. Normal values of FIE have a negative effect on the relative risk, increasing it by a factor of 7.63 for each unit increase of FIE. Very low values of FDE and FIE are both significant in term of increasing the risk of failure.

FDWC has a negative effect on the risk of failure, increasing it by a factor of 9.28 for each increase of FDWC in the normal range, and more than 3 times for high values. No estimate for low values of FDWC has been obtained since there seem to be no value of FDWC that appears to be to small in the dataset considered so the dummy variable FDWCA has not been used (see Table 1).

Analogous results have been obtained considering the Cox model for age (see Table IV). All the ratios show almost the same significance of Table III apart from ROS that is not significant at all, even for very low values. This can be explained considering that these ratio considers the performance of the firm in the previous year and its importance is referred to a limited time interval. On the contrary, the significance of the financial ratios is explained by the fact that they are mainly structural data referring to long term resources of the companies.

ESTIMATES OF THE COX-MODEL FOR YEAR AS TIME TO EVENT VARIABLE

Ratios	Cox Model - year		
	A	NORMAL	B
ROE	ns	ns	ns
ROI	ns	ns	ns
ROS	2.07(*)	ns	ns
CT	ns	ns	ns
ES	ns	0.011(*)	7.43(***)
FDE	4.98(***)	ns	ns
FIE	4.46(**)	7.63(***)	ns
ST1	2.92(*)	ns	ns
ST2	ns	ns	0.28(**)
WCC	ns	ns	ns
FDWC	--	9.28(**)	3.31(**)

(*)=significant at 5%; (**)=significant at 1%; (***)=significant at 0.1%

ESTIMATES OF COX-MODEL FOR AGE AS TIME TO EVENT VARIABLE

Ratios	Cox Model - age		
	A	NORMAL	B
ROE	strata	ns	ns
ROI	ns	ns	ns
ROS	ns	ns	ns
CT	ns	ns	ns
ES	ns	0.01(**)	ns
FDE	5.96(***)	ns	2.27(*)
FIE	3.57(**)	6.19(**)	ns
ST1	3.84(***)	ns	ns
ST2	ns	ns	0.39(*)
WCC	ns	ns	ns
FDWC		9.28(**)	3.31(**)

(*)=significant at 5%; (**)=significant at 1%; (***)=significant at 0.1%

In all the models fitted the size has been a significant variabile to explain the variations of hazards as it was also evident from the Kaplan-Meier estimates in Figures 3 and 4. Namely, small and medium companies have a risk of failure that is always about 60-70% that of micro companies

Conclusion

We performed a longitudinal and a cross sectional analysis based on a sample of 100 Italian non listed companies out of which 50 are bankrupted and 50 are still active on the market over the period 2000 - 2010.

The results of both analyses show that in Italy economic ratios, relying on financial statement data and based on estimations, such as ROE, are not significant in predicting companies failure.

On the contrary, financial ratios and key performance indicators not affected by estimations, such as ebitda/sales, are significant in predicting companies bankruptcy.

This could be explained by the fact that Italian companies tend to minimize net income and the other economic margins based on estimation for tax purpose. This means that the economic ratios, calculated through financial statement, do not reflect the real company's performance and do not measure his real health.

On the contrary, financial ratios and items connected to solidity are much more important and more accurate in predicting companies failure because they are an important indicators for getting loans and other financial means necessary for firms' survival.

References:

Altman E.I, "Financial Ratios, discriminant analysis and prediction of corporate bankruptcy", Journal of finance, 23(4), 1968, 589-609.

Bandopadhyaya, A., "An estimation of the hazard rate of firms under Chapter 11 protection", The Review of Economics and Statistics 76, 1994, 346-350.

Beaver W.H., "Financial Ratios as predicots of Failure", Journal of Accounting Research 4, 1966, 71-111.

Bednarski, T., "On sensitivity of Cox's estimator" Statistics and Decisions, 7, 1989, 215-228

Boardman, C. M., J. W. Bartley, R. B. Ratliff. "Small business growth characteristics", Amer. J. Small Bus. 5(Winter), 1981, 33–42.

Breiman L., Friedman J. H., Olshen R. A., and Stone, C. J., Classification and Regression Trees. 1984, Wadsworth.

Breslow, N.E., "Covariance analysis of censored survival data" Biometrics, 30, 1974, 89-99

Cain, K.C., Lange, N.T., "Approximate case influence for the proportional hazards regression model with censored data", Biometrics, 40, 1984, 493-499

Chancharat, N., Davy, P., McCrae, M., and Tian, G., "Firms in financial distress, a survivaldmo el analysis" School of Accounting and finance, University of Wollongong, 2007

Chen, K. C. and Lee, C. J., "Financial ratios and corporate endurance: A case of the oil and gas industry", Contemporary Accounting Research 9(2), 1993, 667-694.

Coats, P.K and Fant, L.F., "A neural network approach to forecasting financial distress", The Journal of Business Forecasting, 1992, 9-12.

Cox, D. R., "Regression and life tables", Journal of the Royal Statistical Society 34, 1972, 187-220.

Crapp, H. and Stevenson, M., "Development of a method to assess the relevant variables and the probability of financial distress", Australian Journal of Management 12(2), 1987, 221-236.

Efron, B., "The efficiency of Cox's likelihood function for censored data", JASA, 72, 1977, 557-565

Frydman H., Altman E.I., Kao D.L., "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress", The Journal of Finance, Vol. XL, N. 1. 1985, 269-291.

Hastie, T., Tibshirani, R., Friedman, J., The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Springer, New York, 2001.

Huarng, K., Yu, H.K., and Chen, C.J., "The application of decision trees to forecast financial distressed companies", In 2005 International Conference on Intelligent Technologies and Applied Statistics Proceeding, Taipei, Taiwan, 2005

Kaplan, E. L.; Meier, P., "Nonparametric estimation from incomplete observations", J. Amer. Statist. Assn. 53, 1958,457–481

Kim, M.H. and Partington, G. The dynamic prediction of company failure. The 21st Australasian Finance and Banking Conference, pp. 41-47. Sydney, 2008.

Laitinen, E.K. and Luoma, M., "Survival analysis as a tool for company failure prediction", International Research Journal of Finance and Economics, 16 2008, 673-678

Lane, W. R., Looney S. W. and Wansley, J. W., "An application of the Cox proportional hazards model to bank failure", Journal of Banking and Finance 10(4), 1986, 511-531

Nieddu, L. and Vitiello, C. "Determinants of the duration of contracts: a longitudinal study in the Italian fishery sector" Italian Review of Economics, Demography and Statistics, LXVII(1), 2013, 135-142

Odom, M. D. and Sharda, R. "A neural network model for bankruptcy prediction", Proceedings of the IEEE International Joint Conference on Neural Networks *(IJCNN)* 2, 1990, 163-168.

Ohlson J.A., "Financial Ratios and probabilistic prediction of bankrucy", Journal of Accounting Research, 18(1), 1980, 109-131.

Romer,A., "A comparative analysis of determinats of financial distress in French, Italian and Spanish firms" Working paper, Danmarks Nationalbank, Copenhagen, 2005.

Shumway T. "Forecasting bankruptcy more accurately: A simple hzard model", Journal of Business 74, 101-124.

Tam, K. Y. and Kiang, M. Y., "Managerial applications of neural networks: The case of bank failure predictions", Management Science 38, 1992, 926-947.

Therneau T. M., Grambsch P. M., Modeling Survival Data, Extending the Cox Model, Springer, 2000.

Wilson, R. L. and Sharda, R., "Bankruptcy prediction using neural networks", Decision Support Systems 11, 1994, 545-557.

Zavgren, C.V., "Assessing the Vulnerability to Faiulure of American Industrial Firms: a logistic analysis", Journal of Business Finance and Accounting, 12, 1985, 19-46