Modelling default risk: systematic literature review and model enhancements through event data, personal risk propensities and closeness to local communities

PhD Thesis

PhD program in Economics and Finance

La Sapienza University of Rome

Supervisor: Prof. Gianluca Vagnani

Author: Francesco Mazzurco

Student number: 1488775

Date: 27th September 2024

ACKNOWLEDGEMENTS

I would like to thank Professor Gianluca Vagnani and Professor Rita Laura D’Ecclesia of La Sapienza University of Rome for their support, comments and feedback throughout my PhD program and thesis writing. Furthermore, I would like to thank my family and friends for their support, interest and patience during these years of studying and research.

ABSTRACT

The main focus of this thesis is to provide a comprehensive assessment of default risk modelling: the relevant literature, the empirical implementations and the levers driving improvements in prediction's accuracy. The research begins with an overview of the main features of such models and the analysis of the relevant theoretical framework through a systematic literature review. The first chapter's contribution to the research field provides an updated comprehensive assessment of the existing literature and identifies best practices and choices to optimize model accuracy. The following chapter lays the ground for the empirical implementation of the models to validate the applicability of the results of the systematic literature review and select the *base model* with the highest accuracy. In the third chapter, different enhancements are tested on the *base model* to build a new integrated model. The enhancements leverage on complementary research fields (i.e. behavioral finance) and aim to feed the model additional layers of soft information, complementing the core evidence derived from accounting and financial data.

INTRODUCTION

The analysis of default risk has been a key regulatory focus in recent years, especially in EU peripheral countries, due to particularly high levels of defaulted – *non-performing* – exposures within the main financial institutions. As of the end of 2023, non-performing loans (NPL) ratios have materially improved across the EU, dropping on average below 2% of total loans. The improvement has been mainly driven by decisive and targeted action plans executed in southern countries, for example Italy and Greece, where NPL ratios were in the double-digits area. Despite the mentioned developments, the impact of the deteriorated macroeconomic environment, in particular due to the post-pandemic and high inflation related effects, is still uncertain on the economies. All these factors drive the relevance of the research topic in the present economic and financial environment.

Moreover, COVID-19 wide stimulus packages, supported both at the national and EU level have potentially resulted in skewed economic and financial performance indicators. This can result in material modelling errors in default prediction since accounting ratios represent the core of the input information for the analysis. In the thesis, the applicability of additional theoretical structures – i.e.

borrowing from the behavioral finance literature – is presented and tested, with the goal of offsetting potential shortfalls of accounting and financial ratios, in order to reduce the information lag and allow for a more accurate valuation of the company's performance.

In the first chapter, the key inputs and features of a default-prediction model are presented and will be mapped as part of the systematic literature review. The models are generally built on the basis of a set of economic and financial information of companies throughout several years. Therefore, the first key decision of a scholar approaching the subject is the limitation of the scope of the research. For example, this can be related to the type of companies in the database, i.e. listed firms or SME, or to a specific geographic location. The next step is to decide how to select the relevant sample to achieve the best prediction results, i.e. the accuracy of the model to predict whether a company will default or not. Sample selection can be defined as matching, i.e. when the input data fed to the model is split equally between defaulted and non-defaulted companies, or random, i.e. when the database reflects the true default rate of the selected companies.

The remaining key inputs of the model are related to the inclusion, or not, of external indicators and the definition of defaults. The former is extremely important, for example, to feed the model relevant information on the broader economic cycle or qualitative information that can be relevant and indicative of financial distress. These can provide some offset to the inherent lags of economic and financial information, that are usually published yearly or quarterly at best. For what concerns the definition of defaults, representing the dependent variable in the model, this can range from an early situation of financial distress to the various judicial definitions, i.e. bankruptcy. Finally, the main modelling techniques are presented and compared, from the early regression methods to most recent approaches, highlighting the main differences and related upsides and downsides in performance and/or in the interpretation of the results.

The variables and definitions presented in the previous paragraph are the core of the systematic literature review that is performed in the first chapter. Herein, 97 empirical studies have been selected from published journal articles on the basis of a common research string across EBSCO, Science Direct and Google Scholar portals, and the key features of each one have been mapped and compared based on the final accuracy rates: type I, i.e. the ability to correctly identify defaulted firms, and type II, i.e. correct identification of healthy companies. The aim of the analysis is to identify the main levers and combination of inputs that drive the performance of the models and whether any common pattern can be found among the different studies. Results point to statistical significance for the sample selection technique in the overall accuracy of the models and to a broad consensus on the benefits from the inclusion of external indicators and the limitation in scope to only listed firms. Similarly, conditional probability models are the preferred choice among authors due to the upsides connected to the shape of the output and clear interpretation of the impact of each input variable.

In the second chapter, the empirical application of the models is presented, comparing results from conditional probability models, i.e. logit, and neural networks, and a *base model* with the highest prediction accuracy is identified. The training and testing have been performed on the sample of existing companies from the AIDA dataset. This includes 61,675 Italian firms with accounting and financial information related to the 2009-2018 decade, providing a total of 528,474 observations. The Italian landscape is non-homogeneous due to differences in the economic development in the country, i.e. north compared to the south, while at the same time containing a material portion of small and medium-size enterprises (SME). These can be the ideal conditions to test the reliability of prediction models in forecasting defaults of companies of different sizes and in different environments of economic development. Moreover, the chosen observation window from 2009-2018 incorporates many stages of the economic cycle in Italy, from the impact post the financial and European crisis to the following path to recovery. Therefore, this gives the possibility to test the solidity of the model under different macroeconomic contexts. The results are in line with the literature consensus for what concerns the chosen theoretical framework as well as the very solid foundation of the models on the predictive ability of financial ratios calibrated through conditional probability models, i.e. logit regression, which provided the overall higher accuracy rates. The logit model of chapter 2 will constitute the *base model* on top of which the enhancements of chapter 3 will be tested.

In the third chapter, further model iterations are run on a set of additional input variables to test whether these are significant and can increase model accuracy. Three sets of layers of soft information are added to the logit model presented in chapter 2: event data, personal risk propensities and closeness to the center of the local communities. These result in a new integrated model whose performance is tested in the chapter. For event data, we refer to all the qualitative information connected to published news on the company that can be interpreted as an early-warning sign of financial distress. Thanks to a series of routines on online search portals, articles are scraped as simple text and then compared against a set of sensitive keywords generally pointing to a situation of financial distress. The second and third layers are applied to a sub-sample of small and medium-size enterprises (SME) and relate to: (i) additional data linking the information available on the age of the entrepreneur and the firm to potentially higher or lower risk propensities and (ii) to the closeness to the center of the local communities. For the purpose of our research, we have identified those as the closer religious institution. These two layers can represent a way to enrich the original database of input data with relevant information specific to the company and the nearest local economic center. These represent an original contribution of the present work, with a theoretical background imported from the behavioral finance and political economy literature that the present study demonstrates to be also relevant in bankruptcy prediction and enabling improvements in prediction accuracy. Finally, significance has been found for the three layers with positive impact on model's performance. The integrated model results in higher accuracy in the correct identification of defaulted and non-defaulted firms, compared to the *base model*, specifically improving the classification accuracy in SMEs, which, as stated, constitutes a relevant portion of the sample.

Table of Contents

1. SYSTEMATIC LITERATURE REVIEW TO OPTMISE DECISIONAL STEPS IN BUILDING A BANKRUPTCY PREDICTION MODEL

1.1 INTRODUCTION

Default risk modelling has always been widely covered by researchers in consideration of the relevance of the topic both in the academic and in the business world.

From Beaver (Beaver, 1966) and Altman (Altman, 1968) at the end of the 1960s, the scope of research on bankruptcy prediction models has covered not only listed and big corporations, but also small and medium sized enterprises (SME) and retail consumers requesting a form of financing. For enterprises, the importance of financial ratios has been recognized as one of the key factors in the prediction of future financial distress. Most of the studies also highlighted the need for different model's calibrations, for example, when SME are in scope. Listed companies usually have a higher quantity and quality of information due to being public, state regulation, mandatory audit and public scrutiny. Moreover, in the analysis of the default probability, data quality and enhancement through non accounting or financial information have been demonstrated to be a way of increasing the accuracy of the model. Throughout the thesis, we will refer to accuracy of the model as the ability to correctly identify both defaulted and healthy companies. Some studies also showed improvements in model accuracy from the inclusion of macro-economic variables because the importance of the surrounding economic environment can be a key factor in bankruptcy prediction.

There is a considerable body of literature on this theme with great variety of frameworks, from the calibration methodology to the sampling distribution and selection, making it difficult to identify a better overall calibration process. In this thesis we are performing a systematic review of the existing literature to justify the choices made. A quantitative methodology is deployed to select the relevant papers and map the relevant variables to present a comprehensive and standardized view of the different assumptions and techniques from each research and provide the basis for a framework for the development of new models.

The present chapter builds on top of various systematic reviews contributions (Laitinen et al., 2023; Almeida, 2023; Veganzones and Severin, 2021; Calli and Coşkun, 2021; Kim et al., 2020; Farooq and Qamar, 2019; Scherger et al., 2019; Shi and Li, 2019; Agarwal and Patni, 2019; Alaka et al., 2017) and presents the existing literature with the aim of (i) providing an updated contribution to the field, (ii) building a solid foundation of the present research which will be leveraged for the empirical tests in the next chapters, and (iii) proposing an assessment of statistical significance between the chosen input variables and the performance metrics, i.e. type I and type II error rates, of the models.

The chapter is structured with preliminary considerations regarding the existing systematic literature review, which is followed by the overview of the key inputs, and related theoretical backgrounds, needed to build the model. Then, the process of the systematic literature review performed in the thesis is presented, starting with the description of the research string used to select the relevant papers. The key inputs described in *paragraph 1.3* will constitute the variables mapped as part of the systematic review. As a final step, all the collected data will be analyzed with the aim of identifying best practices and main performance levers to optimize the model's accuracy.

1.2 PRELIMINARY CONSIDERATIONS

The listed contributions to the systematic literature review all aim to assess whether a specific combination of input data and modelling technique can result in a better overall prediction ability of the model. The task is not straightforward as different structural impediments are encountered by researchers in the comparison across studies.

In one of the most recent contribution (Laitinen et al., 2023), the authors have tracked all these systematic reviews for a period comprising the last 40 years of research. The identified limitations for the comparison mainly concern to the inherent heterogeneity in the base assumptions used to build the prediction models and can be related, in particular, to the lack of failure theory and empirical limitations of failure studies.

The lack of a standardized failure theory across studies is the main constraint when comparing accuracy results across studies, this is mainly due to the heterogeneity of the default concept itself. Differently than other research fields, where the independent variable can be objectively identified, i.e. in the medical fields where a patient has – or not – a certain condition, in bankruptcy prediction there is a degree of subjectivity for what concerns the definition of default. As better described in the next paragraph, this can range from mild stages of distress, i.e. late payments, to the judicialization procedures activated by the creditors. Understandably, the more severe the stage of distress, the easier it is to correctly predict the default. Moreover, additional challenges in the comparison of studies stem from the empirical nature of the modelling exercise, leading to the selection of predictors which can be sample specific, hence resulting in a model which is also specific to that sample. Finally, different modelling techniques can be coded with different functional forms, adding to the overall complexity of the comparison.

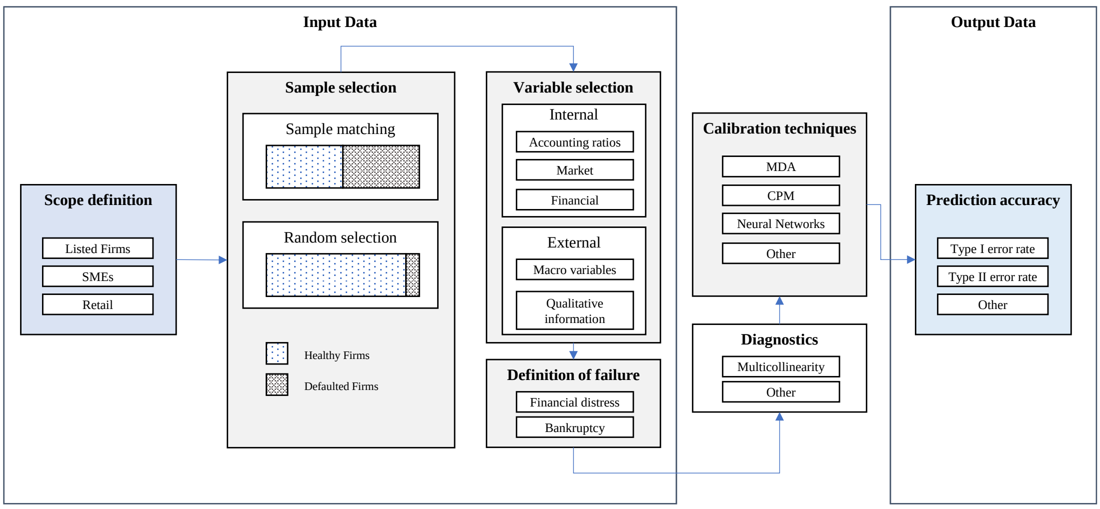
For what concerns the empirical development of the prediction model, the core of the information lays in the financial and accounting data. While some drawbacks may arise from (i) the assumptions regarding the accuracy of the information, especially relevant for smaller companies without audited statements, and (ii) the inherent lag of the information, published annually, these can be partially remediated by feeding additional qualitative information.

In light of the above, the systematic literature review has been performed to track all the relevant variables with the aim of facilitating the comparison across studies. Consequently, the information collected also include data on the key assumptions and decisions made – and disclosed – in each research. For example, these can be related to the choices made for the selection of the sample, the modelling technique, and the type of variables used as input of the model.

1.3 KEY INPUTS OF THE MODEL AND THEORETICAL BACKGROUND

In the calibration process of a default prediction model there are key decisions to be made, starting from the selection of the explanatory variables and their distribution in the training sample to the regression technique. In Figure 1, an overview of the key decision steps necessary for the calibration is presented: from the preparation of the training sample, including the selection of the firms in scope, the relevant variables, and the distribution of defaults to the choice of the regression technique and analysis of the results.

FIGURE 1 Default prediction modelling – Calibration process and key decisions



Scope definition and sample selection

In the complete population of active firms, the defaulted or bankrupt ones represent a small percentage of the overall population. When only bankrupt firms are taken into consideration, the probability of the event is rarely exceeding 1% (Gruszczynski, 2019). This increases if a wider definition of default, that extends to a situation of financial distress, is chosen in the research.

The a-priori distribution in the training sample is one of key choices of the research as this will be impacting the regressors of the chosen modelling technique. The literature is divided between the sample matching technique and the use of an initial sample of random firms. In the former, every one of the selected failing enterprise is matched with a similar healthy company, this structures has been chosen in the first work by Altman (Altman, 1968) among others.

We define $\pi(t), 0 < \pi(t) < 1$ as the failed proportion of failed firms in the complete population in year *t*, and p_t^{π} as the unbiased probability of default in the estimation sample. Both are calculated as a function of a set of explanatory variables *X*. For abbreviation purposes in the following equations, *t* will be omitted in the calculations. Hence, applying the Bayes theorem (Skogsvik and Skogsvik, 2013):

$$(1)\quad p_t^{\pi} = p(defaulted \mid [X])^{\pi} = \frac{\pi * p([X] \mid defaulted)^{\pi}}{\pi * p([X] \mid defaulted)^{\pi} + (1-\pi) * p\big([X] \,\big|\, non_{defaulted}\big)^{\pi}}$$

$$= \left[\frac{\pi * p([X] \mid defaulted)^{\pi} + (1-\pi) * p([X] \mid non\_defaulted)^{\pi}}{\pi * p([X] \mid defaulted)^{\pi}}\right]^{-1}$$

$$= \left[\frac{\pi * p([X] \mid defaulted)^{\pi}}{\pi * p([X] \mid defaulted)^{\pi}} + \frac{(1-\pi) * p([X] \mid non\_defaulted)^{\pi}}{\pi * p([X] \mid defaulted)^{\pi}}\right]^{-1}$$

$$= \left[1 + \frac{(1-\pi)}{\pi}\left[\frac{p([X] \mid non\_defaulted)^{\pi}}{p([X] \mid defaulted)^{\pi}}\right]\right]^{-1}$$

The in-sample probability can be defined with, $\theta(t), 0 < \theta < 1$ and represents the a-priori distribution in the selected sample:

$$(2)\quad p_t^{\theta} = p(defaulted \mid [X])^{\theta} = \frac{\theta * p([X] \mid defaulted)^{\theta}}{\theta * p([X] \mid defaulted)^{\theta} + (1-\theta) * p([X] \mid non\_defaulted)^{\theta}}$$

$$= \left[1 + \frac{(1-\theta)}{\theta}\left[\frac{p([X] \mid non\_defaulted)^{\theta}}{p([X] \mid defaulted)^{\theta}}\right]\right]^{-1}$$

Then, we can define the ratio of defaulting population from Equation 1, assuming that the chosen sample has been randomly drawn from the sub-population of defaulted and non-defaulted firms:

$$(3)\quad \frac{p([X] \mid non\_defaulted)^{\pi}}{p([X] \mid defaulted)^{\pi}} = \frac{\pi}{(1-\pi)}\left[\frac{1}{p(defaulted \mid [X])^{\pi}} - 1\right] = \frac{p([X] \mid non\_defaulted)^{\theta}}{p([X] \mid defaulted)^{\theta}}$$

And, substituting (2) in (3), we obtain:

$$(4)\quad \begin{cases} p_t^{\theta} = \left[1 + \frac{\pi}{(1-\pi)} * \frac{(1-\theta)}{\theta}\left[\frac{1-p_t^{\pi}}{p_t^{\pi}}\right]\right]^{-1}, in\ case\ of: \pi = \theta, this\ reduces\ to\ p_t^{\theta} = p_t^{\pi} \\ p(defaulted \mid [X])^{\pi} = p_t^{\pi}\ AND\ p(defaulted \mid [X])^{\theta} = p_t^{\theta} \end{cases}$$

As shown in Equation 4, the in-sample ratio of defaulted firms can be presented as a function of the true probability of default in the complete population. When the drawn sample is truly representative of the population, Equation 4 reduces to the equality $p_t^{\theta} = p_t^{\pi}$. In case of $\pi < \theta$, the estimations from the model are positively biased and calibrated coefficients are likely to be higher than unbiased regressors.

We now shift our focus on the determination of the cut-off point, under the assumption of using a Conditional Probability Model (CPM), logit belongs to this family of models, in the calibration process. The calibration of the cut-off point is a trade-off process that involves an estimate of misclassification costs. For example, this can be optimized in relation to the overall accuracy of the model or be more sensitive to type I or type II error rates. The former tests the rate of misclassification of failed firms (type I error rate), the latter the rate of misclassification of non-failed firms (type II error rate). We define the error rates as:

$$(5)\quad \begin{cases} (\bar{e}) = \left[\left(e_{type_I}\right) + \left(e_{type_{II}}\right)\right]/2 \\ \{(\bar{e})\}^{\pi} = \pi * \left(e_{type_I}\right) + (1-\pi) * \left(e_{type_{II}}\right) \\ \{(\bar{e})\}^{\theta} = \theta * \left(e_{type_I}\right) + (1-\theta) * \left(e_{type_{II}}\right) \end{cases}$$

From Equation 5, the misclassification cost is:

$$(6)\quad \begin{aligned} & misc(\bar{e}) = w_1 * misc(e_1) + w_2 * misc(e_2) = \pi * \left(e_{type_I}\right) * misc(e_1) + \\ & + (1-\pi) * \left(e_{type_{II}}\right) * misc(e_2) \end{aligned}$$

And the optimal cut-off point is then estimated as:

(7) $IF\ \ p(defaulted \mid [X])^{\pi} * misc(e_1) > (1 - p(defaulted \mid [X])^{\pi}) * misc(e_2)$, THEN firm is classified as defaulted

$$(8)\quad opt_{cut} = 1/\left[1 + \frac{misc(e_1)}{misc(e_2)}\right]$$

From Equations 1 to 4, it is shown how the bias in the initial sample selection is transmitted to the final default prediction of the model. In the case of unweighted misclassification cost function – the cost of misclassified defaulted firm is assumed to be the same as misclassified non-defaulted firm – there is a high chance of higher regression coefficients and a higher cut-off point, resulting in over-classification of defaulted firms (Zmijewski, 1984). More specifically the type I error rate is expected to be more conservative while the type II error rate less precise.

Variable selection

The main issue related to the inclusion of a high number of variables in the model is the presence of multicollinearity in the independent variables. The information and the explanatory power contained in one of the regressor overlaps with one or an entire set of variables. In this case, it becomes more difficult to break down between the unbiased estimate from the model and what is instead the results from a priori inter-relationships between financial or economic ratios.

We consider a linear regression model and define the variance-covariance matrix for the β regressors as:

$$(9) \quad V(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \frac{\sigma^2}{\sum(x_{ij} - \bar{x}_j)^2} * \frac{1}{1 - R_j^2}$$

In Equation 9, R_j^2 is squared correlation factor from the regressor from the variable x_j with the other variables. Therefore, a higher of R_j^2 increases $V(\hat{\beta})$. The ratio $\frac{1}{1-R_j^2}$ is also called variance-inflation factor (VIF) and reflects the degree to which the variance of the regressor increases as consequence of correlation in the regression coefficients – in case of variables with zero correlation among themselves, VIF reaches the maximum value of 1.

In Tinoco and Wilson (Tinoco and Wilson, 2013) to gauge the degree of multicollinearity, VIF tests are computed. Even though there is not a formal criterion to establish a threshold for the tests, a low-test result level is a strong indication of absence of multicollinearity in the model and therefore more stable coefficients.

The variability of the regressor has an impact also on the statistical tests related to the explanatory power of the model and its predictive ability. In bankruptcy prediction modelling, trying to improve the model by adding additional variables, without running any form of collinearity diagnostic process can lead to biased estimators. One way that has been proposed in many of the selected studies in our sample is to move to external indicators, related to the macroeconomic cycle or qualitative traits of the specific firm (audited or not), in order to increase the accuracy of the estimations.

Calibration techniques

The first Altman's model (Altman, 1968) as well as more from others in the following decades were developed using the multi discriminant analysis, MDA, technique. The methodology aims to classify objects into categories based on their observed characteristics, the weights of the discriminant function can be interpreted as the relative importance of each characteristic in the classification process.

In Altman (Altman, 1968) the sample was constructed with 66 corporations, half of which filed bankruptcy, mean asset size was $6.4m (corresponding to approximately $52,4m in 2020 assuming an average annual inflation rate of 3.7%). Five financial ratios are selected covering different categories of a firm's health state: liquidity, profitability, leverage and solvency – this first version of the model, presented in Table 1, is only applicable to listed enterprises because of the market capitalization component.

TABLE 1 Z-Score – Altman (1968)

(1)	$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$
	X_1: Working capital/Total assets X_2: Retained earnings/Total assets X_3: Earnings before interest and taxes/Total assets X_4: Market value equity/Book value of total debt X_5: Sales/Total assets

There are a set of assumptions needed for this type of calibration: independent variables need to be multivariate normal, the covariance matrices of the class of failed and non-failed companies need to be equivalent, prior probabilities of failure are to be taken under consideration and there should not be any degree of multicollinearity. Multicollinearity, as will be discussed later, can be defined as the degree of linear dependency in a multivariate model between the independent variables, the issue is relevant to bankruptcy prediction modelling as some financial ratios might be influenced by similar factors or trends.

The discriminant analysis function is implemented to create a boundary to divide defaulted and healthy firms from the selected sample as a linear combination of the explanatory variables. The final score will be derived from:

$$(10) \quad z_i = \sum_{k=1}^{N} \beta_k x_{i,k} = \beta' x \; where \; \beta = V^{-1}(x_1 - x_2) \; and \; k = 1,2$$

In Equation (10) V represent the variances and covariances matrix of the selected independent variables while $x_1$and x_2 are the means of the defaulted and non-defaulted sub-samples (centroids). The β coefficient aims to maximize the group separation – or distance – between the means.

$$(11) \quad \max_{\beta} \square = \frac{|z_{nd} - z_d|}{\sigma_z} \; which\ becomes \; \frac{(\beta' x_{nd} - \beta' x_d)^2}{\sigma_z^2} = \frac{(\beta' x_{nd} - \beta' x_d)(\beta' x_{nd} - \beta' x_d)}{\beta' V \beta} = f(\beta)$$

In Equation 11, we consider the square of the expression to avoid working with absolute values and compute the gradient of $f(\beta)$.

$$(12) \quad \frac{\partial f}{\partial \beta} = \frac{2(\beta' x_{nd} - \beta' x_d)(x_{nd} - x_d)\beta' V \beta - 2V\beta(\beta' x_{nd} - \beta' x_d)^2}{(\beta' V \beta)^2} = 0$$

Solving Equation 12 we obtain Equation 10. Under the assumption of multivariate normal distribution, the z-value from Equation 10 can be transformed into a probability of default using the below equation. In Equation 13, we have defined π as the prior probability of default and *c* as the chosen cut-off point.

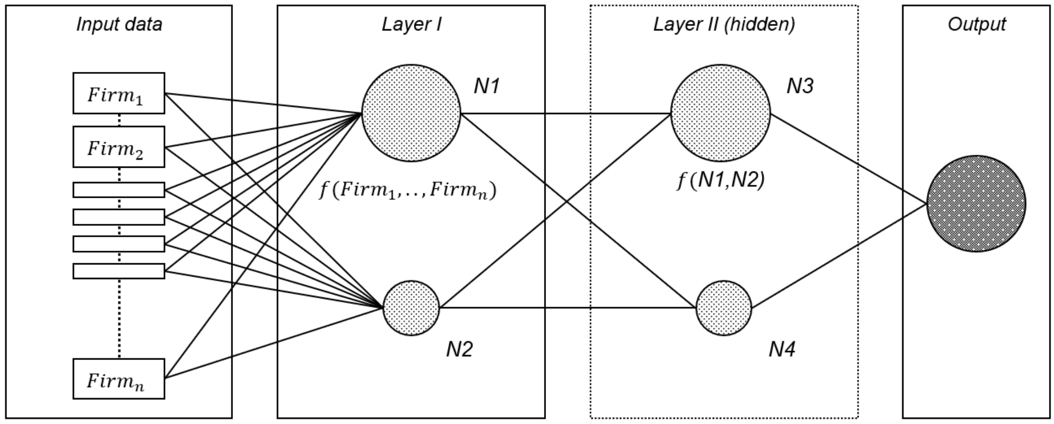
$$(13) \quad p(default\,|x_i) = \left(1 + \frac{1 - \pi}{\pi} e^{z_i - c}\right)^{-1}$$

CPM are an alternative to the MDA technique because of less stringent assumptions and easier output interpretation mainly because of the binary output of the function, defaulted or not defaulted, ranging from 0 to 1. Even though both models are based on the resemblance principle, as firms are categorized to either one of the failed or non-failed group according to which one they are closer to, in CPM no assumption is needed regarding prior distribution of defaults, even though an accurate calibration of the cut-off point will then be necessary, and distribution of the selected variables. CPM models share with MDA a material sensitivity to extreme non-normality (McLeay and Omar, 2000) and multicollinearity. Mathematically, the linear relationship is adjusted through an exponential transformation, the domain of the function is now [0,1].

$$(14) \quad z_i = \left(1 + e^{-c - \sum_{k=1}^{N} \beta_k x_k}\right)^{-1} + \varepsilon_i$$

Moreover, neural networks (NN) technique is also starting to be adopted with as good results as the more traditional models. Calibration techniques that leverage neural networks are based on an iterative mechanism. The network consists of nodes (the neurons), distributed in layers, that are connected to each other by a simple relation (synapses). The first layer of the system receives the initial information, made of the set of explanatory variables selected in the study, and processes this with a linear or non-linear function. Results are transmitted to the next layer. In the example described in Figure 2, the information received by the first layer is then processed by a second layer in the network and a final output is obtained providing the indication of defaulted or non-defaulted firm. Neural networks are often classified as inductive models as results are obtained by simulating the learning process of humans. In the case of default prediction modelling this is the process of identifying the correct weights to be applied at the initial set of inputs and to the intermediate inputs transmitted by the hidden layer. The particular structure of the network has the downside of a more difficult interpretation of the relationship between the final output and the initial set of explanatory variables. In Charitou et al. (Charitou et al., 2004) a NN model was implemented on listed UK enterprises with sample matching technique, and showed ex-ante good results, better than CPM, in the identification of non-failed enterprises (type II error rate).

FIGURE 2 Example of Neural Network

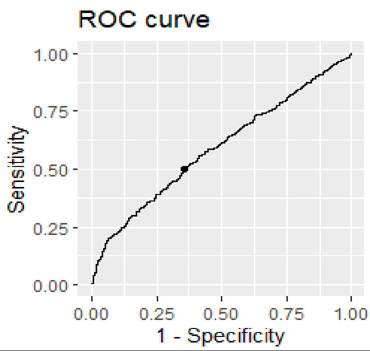


Prediction Accuracy

The performance of a model is generally evaluated with the classification accuracy of the two main status, defaulted or not. Results can be visualized using a confusion matrix, with size $R\ x\ R$. An example commonly found in the literature is presented in Table 2 below.

TABLE 2 Confusion Matrix and ROC curve

	Defaulted	*Non-Defaulted*
Defaulted	$1-\left(\frac{TP}{TP+FN}\right)$	
Non-Defaulted		$1-\left(\frac{TN}{TN+FP}\right)$
Overall accuracy	$1-\left(\frac{TP+TN}{TP+TN+FP+FN}\right)$	



The formulas in the table are a way of computing type I, type II and overall error rates, the variables are defined as follows:

- TP (true positive ratio) is defined as the proportion of defaulted firms correctly classified as defaulted.
- TN (true negative ratio) is defined as the proportion of non-defaulted firms correctly classified as non-defaulted.
- FP (false positive ratio) and FN (false negative ratio) are defined as the proportion of incorrectly classified firms.

The ROC (receiver operating characteristic) curve provides a useful visualization tool to compare model's outputs with different decision thresholds.

1.4 LITERATURE SELECTION AND RELEVANT VARIABLES

Research-string to select the relevant studies

The process of identifying the studies to be collected for the literature review begins with the choice of the database sources containing the appropriate articles, the definition of the research string (Kuizinienė et al., 2022; Shi and Li, 2019) and the extraction of the data and the subsequent cleaning of the acquired database.

The articles were selected among three main sources: EBSCO, Science Direct and Google Scholar, filtered for journal articles while excluding conference papers. The research string was calibrated using the Boolean expression presented in Table 3 and, in consideration of access availability, the time horizon was set from 1995 to 2023.

TABLE 3 Research string

Description	***String***
scope	("bankruptcy" OR "failure" OR "default" OR "financial distress") AND ("score" OR "prediction" OR "rating") AND
target	("companies" OR "company" OR "firms" OR "listed" OR "corporation" OR "corporations") AND
techniques	("MDA" OR "Logit" OR "neural network" OR "regression") AND
statistics	("classification accuracy") AND ("Lachenbruch" OR "validation test")

The string can be broken down in four main sections:

- scope: this is the first filter added to the query to delimitate the perimeter of the analysis to bankruptcy or financial distress default prediction models,
- target: this filter is used to select the scope of the papers,
- techniques: the filter for regression technique has been designed to capture the three main calibration approaches for this type of research: multi discriminant analysis (MDA), conditional probability models (CPM) and neural networks (NN), and a catch-all name – "regression" in the string – to keep and analyze papers where neither one of the most popular regression techniques were used,
- statistics: the filter is used to only extract papers that present a form of statistical validation, in-sample or out of sample, to compare the research frameworks on a common quantitative basis.

The output from the research string has then been analyzed to: *(i)* remove duplicates, mainly due to same research papers selected from both the three sources, *(ii)* filter out the studies not relevant to the research subject, i.e. medical journals (those were included, in most of the cases, because of the word combination of *failure-prediction*) and *(iii)* remove the studies that did not present any quantitative statistical benchmark. The final sample also include relevant studies that were cross-referenced in one or more papers from the quantitative selection process and accuracy results from the models are captured in their standalone form (Alaka et al., 2017).

Coding of the relevant variables and assumptions

The selection of coded variables matches the main themes and assumptions in the calibration process of a bankruptcy prediction model, as described in Figure 1: *(i)* the enterprises in scope of the analysis and the bankruptcy definition, *(ii)* the modelling technique, *(iii)* sample and variable selection and *(iv)* model validation techniques. The four listed buckets have an additional level of granularity as presented in Table 4 below.

TABLE 4 Coded fields across the selected sample of papers

Reference			Sample								Variables							Model							Testing and validation			
Year of publication	Period of the analysis	Authors	Country	Sample size	Bankruptcy % in sample	Sector/Industry	% of industrial companies	% of financial companies	% of SMEs	pre-test on sample	Number of variables	Accounting Ratios	Market ratios	Non-accounting/financial variables	Macroeconomic variables	Binary variables	Transformation of variables	Technique used	Model validation	Model order	Lag	r squared	Default type	Test type	Type I error rate (testing)	Type II error rate (testing)	Type I error rate (validation)	Type II error rate (validation)

Scope of work: Firm's type

The mapped variables include the type of firm, the country where the analysis was performed and the a priori distribution of failed firm in the sample.

The literature on listed firm is the most prominent in the selected research sample, from Altman (Altman, 1968) to Charitou (Charitou et al., 2004) among others. There are some arguments in favor of selecting listed corporations as target of the study. The main ones are related to data availability and trustworthiness as accounting and financial ratios are public information that is usually controlled and audited regularly. This is not the case when the scope of the work is shifted towards SMEs, therefore the need to include additional types of variables to improve the performance of the model. For instance, in Altman, Sabato and Wilson (Altman et al., 2010) the database consisted of approximately 5.8m SMEs non defaulted in the 2000-2007 period and about 66,8k failed enterprises. One feature of the defaulted prediction model was the analysis of non-financial and non-accounting information: the "event data". This information is related to the defaulting on credit agreements or credit payments, the enterprises being late to file their financial statements, or the company being audited or not. While the first indicator might be linked to a level of financial distress, not audited information is assumed to be less reliable and auditors are supposed to prevent a technically insolvent firm from continuing its operations as normal.

The country of risk has been mapped as well in the selected sample of articles: most of the works have been published on American corporations, but also papers where firms were from the United Kingdom, Europe and Asia have been included.

Variable selection, sampling and bankruptcy definition

There is no consensus on the appropriate number of variables to be selected as well as on the accounting ratio category (i.e. liquidity, profitability or coverage) which can be demonstrated to be a better predictor. In Charitou (Charitou et al., 2004), Altman and Sabato (Altman and Sabato, 2007) among others, a stepwise regression is performed to select with an iterative process the variables that give the model the highest accuracy. In other articles, for example Pindado (Pindado and Rodriguez, 2004) and Balcaen (Balcaen and Ooghe, 2006) among others, it is stated that, in consideration of the high correlation among accounting variables, some even share the same denominator, it is possible to get good results even when less ratios are used mainly because the predictive power of the additional predictors has already been accounted for by the existing inputs of the model (Hand, 2004).

The accuracy of the model can be improved thanks to the inclusion of non-accounting information, as mentioned in Sabato (Altman and Sabato, 2007), and macroeconomic data in the model calibration phase. In Liou and Smith (Liou and Smith, 2007), the authors showed how the general context in which the company operates can be a major factor in establishing the probability of default. In this research, six indicators were tested: the GDP, the Industrial Production Index (IPI), the UK base rate, the Producer Price Index (PPI – used as a proxy for inflation), the Retail Price Index (RPI – as proxy for the cost of living) and FTSE All share (to gauge the overall economic trend). Results indicate that the use of macroeconomic variables as a second stage filter can help in the reduction of type II error rate.

Regarding the representation of the defaulted population in the training sample, there is almost equal distribution between matching and random sampling in the selected group of articles. On the modelling technique side, CPM is the most popular choice among the extracted papers, the upsides of using this calibration type have been presented in the previous section.

On the definition of bankruptcy there is not uniformity across the selected sample of articles. In Laitinen (Laitinen, 2010), the author is using a definition of financial distress as the inability of the enterprise to meet its financial obligations when they became due. An argument in favor of this choice is that default prediction models are often implemented with the goal of building an early warning system that still enables the bank/creditor to act at an earlier stage of distress of the client. On the other hand, in Charitou (Charitou et al., 2004), the juridical definition of bankruptcy is preferred as it provided an objective criterion for the classification of firms.

The breakdown of the described results of the literature selected as part of the process are presented in Table 5.

TABLE 5 Literature selection – Summary

	Summary			**Modelling Technique**				**Training**		**Validation**	
	Avg (*)	**Median**	**std dev**	**MDA**	**CPM**	**NN**	**Other**	**Type I**	**Type II**	**Type I**	**Type II**
Sample Size (n)	**225,430**	**284**	**1.06E+06**	**579**	**447,511**	**159,311**	**29,626**				
Sampling technique (b)											
*matching sampling**	53			12	17	22	2	87%	85%	86%	80%
*random sampling**	44			6	21	9	8	81%	85%	78%	82%
Type of firm (b)											
*Listed**	39			9	15	11	4	88%	88%	88%	83%
*Other**	58			9	23	20	6	82%	83%	79%	80%
Number of variables (n)	**19**	**15**	**13**	**18**	**17**	**21**	**27**				
External indicators (b)											
*included**	49			10	18	15	6	83%	83%	81%	82%
*not included**	48			8	20	16	4	86%	88%	83%	80%
Modelling technique (m)											
*MDA**	18							84%	83%	83%	75%
*CPM**	38							80%	83%	77%	81%
*NN**	31							88%	88%	88%	84%
*Other**	10										

** if average is not applicable, in-sample presence is reported*

(b) = binary variable, (n) = actual number from coded variable, (m) = value base on mapping

1.5 FINDINGS FROM THE SYSTEMATIC LITERATURE REVIEW

The relevant assumptions and outputs from each research have been consolidated for the purpose of integrating the findings. One of the methods to gauge the impact of variables is the ordinary least squares multiple regression model. We assume the dependent variable vector to be the measure of accuracy in the bankruptcy prediction models and the matrix of explanatory variable to contain each study characteristics.

The output of two types of regression models is analyzed: for the first model, tests are performed on the type I error rate, on type II error rate for the second model. The structures of the models and results are presented in Table 6 and Table 7.

The input variables that are considered have been selected as they are deemed the most relevant in the determination of the final accuracy of the model:

(i) the sample size – used as control variable,

(ii) the sampling technique – input as binary variable, 1 if the bankrupt percentage of the population is half of the sample, 0 if a random sample population is selected,

(iii) the type of firm – mapped as binary variable, 1 if listed enterprises are in scope and 0 otherwise,

(iv) the total numbers of variables selected in the calibration process of the model – used as control variable,

(v) the presence of external variables included in the calibration process – binary variable, 1 if only accounting ratios have been selected, 0 otherwise,

(vi) the modelling technique – variables has been mapped 0 for univariate models, 1 for MDA, 2 for CPM, 3 for NN and 4 for others.

TABLE 6 Regression models' structure

(I)	$Y_{type\ I} = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$
(II)	$Y_{type\ II} = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \varepsilon$

X_1: Sample size (control variable)
X_2: Sampling technique (binary variable)
X_3: Type of firm
X_4: Number of variables (control variable)
X_5: External indicators (binary variable)
X_6: Modelling technique

TABLE 7 Regression models' results on training and validation samples

	Model I (Type I error rate)			Model II (Type II error rate)		
Variable	Coefficients	Standard Error	t Stat	Coefficients	Standard Error	t Stat
Sampling technique (b)	0.09** (0.08)***	0.04 (0.02)	2.47 (3.23)	-0.08** (-0.003)	0.03 (0.03)	-2.38 (-0.1)
Type of firm (b)	0.11*** (0.08)***	0.04 (0.03)	3.15 (2.99)	0.02 (0.03)	0.03 (0.03)	0.66 (1.26)
External indicators (b)	0.01 (-0.01)	0.03 (0.02)	0.18 (-0.28)	0.06* (-0.05)*	0.03 (0.03)	1.95 (-1.94)
Modelling technique (m)	0.03 (0.03)***	0.02 (0.01)	1.50 (2.3)	0.05*** (0.03)**	0.02 (0.01)	2.87 (2.2)

(b) = binary variable, (n) = actual number from coded variable, (m) = value base on mapping
Training accuracy rate in parentheses. $* p < 0.10$; $** p < 0.05$; $*** p < 0.01$

Analysis of the results

Statistical significance has been observed in the regressor coefficient for sampling technique under Model I and Model II: the initial assumption on the sample selection is statistically relevant in the articles' population. As described in the previous section, the matching-sample technique, each failed firm is matched with a similar non-failed enterprise, can lead to an increase in type II error rate. Results are therefore consistent with the findings in Balcaen and Ooghe (Balcaen and Ooghe, 2006) among others and the theoretical background presented, of more biased parameters when the estimation sample is non-random and that such assumption might result in a more accurate classification of failed firms, but a higher error rate for non-failed firms.

The modelling technique has been cross tested separately to verify whether a particular calibration process can be found to perform better than the others. The sample of papers includes approximately more than two-thirds of entries under CPM and NN, while MDA and other techniques (hazard model and univariate) are only a residual part of the sample. The majority of presence of CPM models might indicate a form of preference for this type of calibration technique, which is also justified by the less stringent assumptions and easier interpretation of the result output compared to the other alternatives. Even if different model accuracy rates have been coded, for example when in a selected research two or more calibration techniques were compared, there might also be an implicit bias in the regressor itself. Research papers where a given modelling technique accuracy rates were not deemed satisfactory would not have reached the publication phase.

Additionally, type of firm regressors have been found significant in Model I. When listed firms are in scope of the research, accuracy rates are higher in our sample. As mentioned in the previous sections, the quality and availability of data for listed enterprises is generally better than SMEs mainly due to higher requirements in reporting standards.

The external indicator variable is borderline statistically significant in the tests even if the type of variables represents a key choice in the framework definition. The accounting variables flag is coded as “1” if, in the calibration process, authors have included market ratios, macroeconomic indicators or other variables. In the selected sample of articles there is, though, a broad consensus on the benefits from the addition of external indicator’s due to their predictive ability of financial distress. External indicators have also the advantage of showing limited collinearity with financial and accounting ratios allowing to incorporate additional information in the model while keeping low the probability of additional bias in the regressors.

TABLE 8 Hypothesis and final results

Variable	***Hypothesis***	***Statistical significance***	***Broad consensus in selected sample***	***Not consistent evidence of from the analysis***
Sampling technique (b)	The chosen sampling technique has a direct impact on the regression coefficients, impacting model accuracy	[x]		
Type of firm (b)	The type of firm has a direct impact on model accuracy, information for Listed Corporation is subject to regular audit, reducing data quality issue in the calibration		[x]	
Inclusion of external indicators (b)	The inclusion of external indicators increases model accuracy		[x]	
Modelling technique (m)	CPM modelling technique, given less stringent assumptions and easier output interpretation, is the preferred choice compared to the other alternatives		[x]	

(b) = binary variable, (n) = actual number from coded variable, (m) = value base on mapping

1.6 CONCLUSIONS

The literature on bankruptcy prediction models is very diverse in scope, assumptions and presentation of results. Using a quantitative approach for both the selection of the sample of studies and analysis of the research findings, we aim to better structure the whole process. There are many key steps for building up a solid model. Starting from the scope definition, the sample and variable selection and the calibration technique. All the steps are obviously relevant and impacting the prediction accuracy of the model. With a systematic review approach, we presented a statistical background that provides the reader with transparency on the framework definition process. Similarly as other studies on the topic, this chapter has some limitations. The results from the selected literature have not been reperformed, this is mainly due to inherent challenges of the subject, i.e. availability on the database used for testing.

Nonetheless, research findings from the quantitative process are consistent with the relevant theoretical background, especially on sample matching and unweighted cut off point resulting in an overall less accuracy in identifying non-defaulted firms. On the other hand, when listed firms are the scope of the research, accuracy rates are higher on average. Regarding the inclusion of external indicators as independent variables, statistical significance has not been found in the regressors, but there is wide consensus in the literature on upsides of these variables: those feed additional information with limited risk of multicollinearity with firm’s internal accounting ratios.

2. FRAMEWORK VALIDATION

2.1 INTRODUCTION AND RELEVANT LITERATURE

While, historically, the first models have been developed using the multi-discriminant analysis technique, MDA, i.e. Altman (Altman, 1968), most recent contributions have been leveraging on conditional probability models, CPM (i.e. logit) and neural network models. Consequently, the models tested in the thesis will belong to these two families. Logit models represent a step forward from MDA with (i) less stringent assumptions in connection to independent variables' distribution and (ii) equal defaulted / non defaulted dispersion matrices (McLeay and Omar, 2000). Moreover, the coefficients of the regressions and their significance can be interpreted as the marginal contribution of each factor to the overall default prediction. At the same time, the function domain ranging from 0 to 1, is particularly useful when estimating a probability of default. Neural network is the other modelling technique tested as performance-wise very similar to CPM model. The main downside of using such models is mostly related to the limited ability of the researcher to trace back the first order variable impact to the overall default probability. Due to these reasons, our tests will aim to assess whether logit model can fit the purpose of providing adequate prediction accuracies on the sample of Italian firms, compared to the other alternative represented by neural network models.

Secondly, the analysis aims to assess the adequateness of the financial and accounting data in providing information of the degree of financial distress for the Italian landscape. Since the very first literature contributions (Beaver, 1966) the scope of indicators has included profitability indicators (i.e., net profits to assets and debt to assets) and liquidity, or financial, ratios (i.e., current ratio, operating cash flows to debt and net working capital). These areas have then evolved to also include indicators flagging the company's leverage. While accounting ratios generally represent the common ground in most of the literature, other authors have also proposed the inclusion of macroeconomic indicators, for example connected to the national or regional gross domestic product. Considering the sample of Italian companies and the national industry mostly composed of SME companies, the selection of financial variables related to the three mentioned macro areas (financial, leverage and profitability) is strongly supported in the wide bankruptcy modelling literature, *leveraging on the analysis prepared in Ciampi et al* (Ciampi et al., 2021; Altman et al., 2020; Baixauli and Modica-Miko, 2010; Calabrese et al., 2016; Castillo et al., 2018; Ciampi and Gordini, 2013; Gupta et al., 2014; Yazdanfar, 2011).

The present study is structured with the first section describing the database and its peculiarities, especially those related to the sample composition and the underlying driving factors as well as the definition of bankruptcy being used in this research. Then, different model's calibration techniques will be tested: logit, stepwise logit and neural network. The model with the highest overall accuracy will be the *base model* used for the analysis of chapter 3.

2.2 DATASET

Geographical distribution

The database consists of ~53.000 Italian firms that submitted accounts information in the 2009-2018 decade, with a total of ~528.000 observations. Of those, defaulted companies represent on average 5% of the population throughout the observation window, the geographical distribution of the dataset is presented in Figure 3. The accounting and financial data as well as the status of the company have been sourced from AIDA.

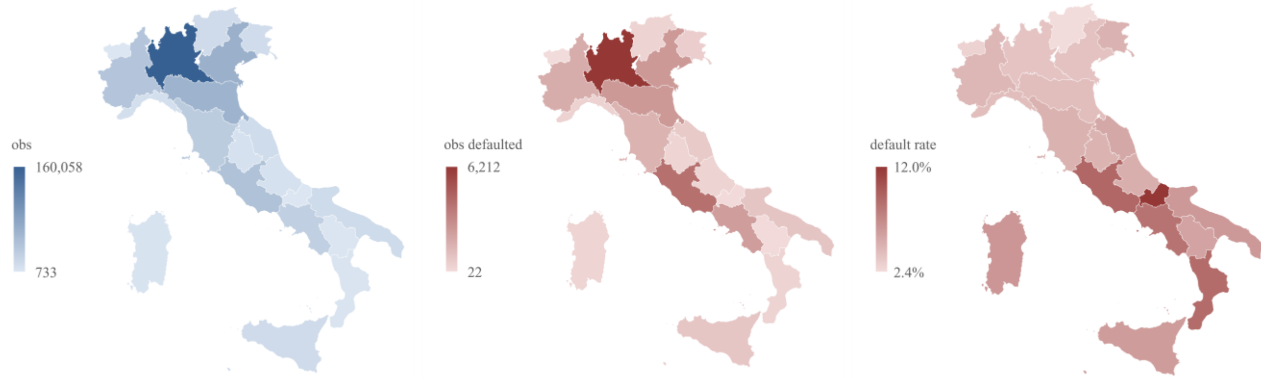
The non-homogeneity of the sample across the different Italian Regions is mainly explained by historically different level and dynamism of economic development. Northern Regions are therefore more represented compared to the Center and the South and present a structurally lower default rate.

Historically, the north-south divide can be traced back to the second half of the 19^{th} century, when the Kingdom of Italy was proclaimed (1861). During these years, the dynamic of the Italian industrialization process concentrated in the north – in the "industrial triangle" specifically – with southern regions lagging on the back of lower education levels, life expectancy and infrastructure development alongside an economic activity mainly focused on extensive agriculture. In the early twenties of the 20^{th} century, up until the war, the gap continued to increase – with GDP average growth of 0.5%, compared to almost 2% in the North.

In the reconstruction years, also known as the Italian *golden age*, the trend briefly inverted toward a convergence due to the strong investments from the government. The GDP average growth was at 5.8% in the South, 1.5% higher than the North, and was strongly linked to the initiatives backed by *Cassa del Mezzogiorno* which benefited from the international economic environment as well as the support from the World Bank and other international organizations.

This did not last long as following events up to more recent years – from Bretton Woods and the oil crisis in the seventies to the more recent globalization trends and financial crisis – restored the divergence trend between the two areas. The latest economic downturns, during the 2008 financial crisis and 2011 European crisis, negatively impacted both areas, but Northern GDP drop was only half of the South which was also affected by the international globalization trends that led to a substantial resizing of its manufacturing structure.

FIGURE 3 Geographical distribution of the observations for the complete population, defaulted companies and default rate



Definition of default

In the Italian regulatory and judiciary system, different degrees of severity are identified in connection with company financial distress. The first warning signals begin as soon as the debtor does not meet the agreed repayment timeline or if there is a reasonable concern about future soundness of the economic and financial business plan.

There are three main buckets provided by the Italian law for the economic reasons that may lead to insolvency proceedings:

- state of crisis: this is defined as "the state of economic and financial difficulty which makes the debtor's insolvency probable, and which, for companies, is demonstrated by the inadequacy of the prospective cash flows to regularly meet the planned obligations[1]".
- state of insolvency, this is triggered when a company defaults on its debt or any other external factor or event prove that the debtor is no longer able to fulfill its obligations. The mentioned state is the first requirement to start a bankruptcy proceeding.
- over-indebtedness, which is the state of crisis or insolvency for consumers and any other entrepreneurs not subject to judicial liquidation.

Therefore, the situation of financial distress represents a wide definition of default which contains the judicial definition of bankruptcy. Bankruptcy is a procedure to liquidate the insolvent

[1] Art. 2 Corporate Crisis Code and Legislative Decree 12th January 2019.

company’s assets with consequent distribution of its proceeding to the creditors, under the supervision of the Court. Alongside bankruptcy, “concordato preventivo” is also a court-supervised judicial procedure in which the debtor can submit a plan of restructuring or discharge of debt in order to avoid bankruptcy. In this case, the debtor must usually meet law requirements of granting fulfillment of at least 20% of the unsecured exposures. Moreover, for a debt restructuring arrangement to be valid this *(i)* must be agreed by at least 60% of the creditors owning the nominal value of the debt and *(ii)* must be validated and approved by an independent expert to gauge reasonableness, feasibility and suitability of the restructuring plan.

In this thesis the definition of default being applied includes *(i)* entry into administration, *(ii)* entry into liquidation, *(iii)* dissolution of the company, *(iv)* entry into debt restructuring agreement, a more detailed list of all the status that led to the default classification is presented in Table 9.

TABLE 9 Company status: in-sample classification distribution

Non-defaulted	**502,167**
Attiva	454,365
Altre cause	23
Cancellazione dal registro delle imprese	33
Cessata d'ufficio perché già iscritta nel registro ditte e non transitata nel registro imprese	77
Cessazione d'ufficio	20
Cessazione d'ufficio su segnalazione registro imprese della sede legale	4
Chiusura del fallimento	3
Chiusura dell'unità locale	223
Conferimento	8
Fusione mediante costituzione di nuova società	737
Fusione mediante incorporazione in altra società	37,355
Liquidazione volontaria	745
Motivo non precisato	10
Ordinanza presidenziale	3
Provvedimento di cancellazione dal registro delle imprese	14
Scissione	673
Sequestro conservativo di quote	10
Trasferimento in altra provincia	7,345
Trasferimento sede all'estero	401
Trasformazione di natura giuridica	17
Trasformazione in sede legale	101
Defaulted	**26,307**
Accordo di ristrutturazione dei debiti	459
Altre cause	20
Amministrazione giudiziaria	57
Amministrazione straordinaria	657
Cancellata d'ufficio ai sensi art. 2490 c.c.	137
Cancellata d'ufficio ai sensi dpr 23/7/2004 n.247	13
Cancellazione dal registro delle imprese	2,929
Cancellazione d'ufficio a seguito istituzione cciaa di Monza	80
Cessazione delle attività nella provincia	50
Cessazione di ogni attività	50
Cessazione d'ufficio	52
Chiusura del fallimento	1,750
Chiusura della liquidazione	784
Chiusura per fallimento	13
Chiusura per fallimento o liquidazione	12
Chiusura per liquidazione	10
Concordato fallimentare	155
Concordato preventivo	1,034
Decreto cancellazione tribunale	9
Fallimento	14,634
Liquidazione	238
Liquidazione coatta amministrativa	224
Liquidazione giudiziaria	38
Liquidazione volontaria	172
Mancata ricostituzione della pluralità dei soci	4
Provvedimento di cancellazione dal registro delle imprese	781
Scioglimento	200
Scioglimento e liquidazione	981
Scioglimento e messa in liquidazione	2
Scioglimento per atto dell'autorità	24
Scissione	8
Stato di insolvenza	600
Trasferimento in altra provincia (Stato di insolvenza, fallita, in liquidazione)	130
Total observations	**528,474**

Accounting and financial ratios

Accounting ratios are the founding pillar of the bankruptcy prediction model and have been collected to provide indication of health of three main areas: liquidity, profitability and leverage – Table 10.

Liquidity (“quick”) and current ratios are both included in the input dataset, those gauge the ability of the company to pay-off short term liabilities using short term assets, the two indicators differ for the more conservative approach being followed for the quick ratio. Lack of liquid asset is often the main cause of financial distress therefore it is expected that the defaulted population would present lower levels of the ratios.

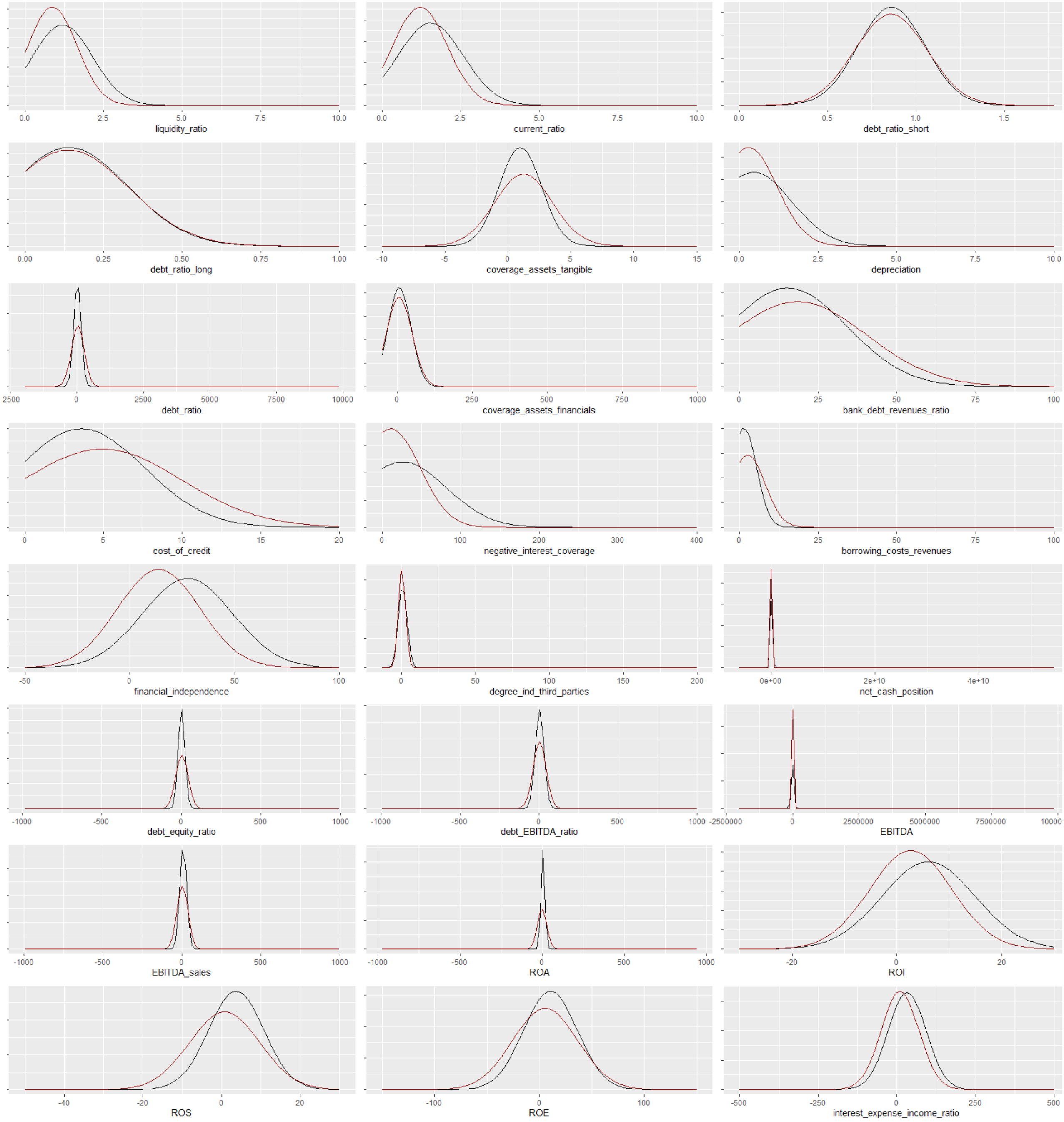
Moreover, level of coverage of medium and longer term liabilities are included as well as indicators of the cumulative profitability of the company.

The distribution of each ratio in the sample and in the defaulted population is presented in Table 11, in most of the indicators the normalized distribution is coherent with the assumption of strong relationship between key accounting and financial indicators and company ability to meet its debt obligations.

TABLE 10 Selected accounting and financial indicators

Area	Indicator	Calculation	Description
Financial	liquidity ratio	(cash + cash equivalents + current receivables + short term investments) / current liabilities	level of coverage of short-term liabilities with very liquid assets (<90 days liquidity horizon)
Financial	current ratio	current assets / current liabilities	level of coverage of current short-term liabilities using short term assets (<1y liquidity horizon)
Financial	fixed-asset (tangible) coverage ratio	((Assets – Intangible Assets) – (Current Liabilities – Short-term Debt)) / Total Debt	level of coverage of total debt with tangible assets
Financial	depreciation	accumulated depreciation / fixed asset	measure of age, value and remaining usefulness of fixed assets
Financial	fixed-asset (financial) coverage ratio	((Assets (Financials) – (Current Liabilities – Short-term Debt)) / Total Debt	level of coverage of total debt with financial assets
Financial	bank debt/revenues ratio	bank debt / revenues ratio	-
Financial	cost of credit	-	-
Financial	net financial position	-	level of coverage of current interest payment with available earnings
Leverage	short-term debt ratio	short term debt / total debt	level of coverage of current interest payment with revenues
Leverage	long-term debt ratio	long term debt / total debt	sum of bank borrowings, short, medium and long-term borrowings, net of cash held in hand and at bank.
Leverage	debt ratio	total debt / total asset	-
Leverage	equity ratio	total equity / total asset	-
Leverage	equity ratio (from external stakeholders)	-	leverage used by the company
Leverage	debt/equity ratio	total liabilities / total shareholders’ equity	leverage used by the company
Leverage	debt/EBITDA ratio	debt / EBITDA	-
Profitability	interest coverage ratio	EBIT / interest expense	leverage used by the company
Profitability	interest/revenues ratio	interest expense / revenues	-
Profitability	EBITDA	-	-
Profitability	EBITDA/Sales ratio	-	-
Profitability	ROA	net income / total assets	earnings generated from invested capital
Profitability	ROI	EBITDA / total assets	-
Profitability	ROS	EBITDA / revenues	-
Profitability	ROE	net income / average shareholder equity	-
Profitability	non-core losses/gains	-	-

TABLE 11 Accounting ratios – in-sample distribution (normalized), defaulted population in red



Macroeconomic environment and indicators

The historical window covers the years from 2009 to 2018. In this time frame the Italian economy received two major shocks: from the US financial crisis of 2008 and the closer and stronger impact of the EU crisis of 2011.

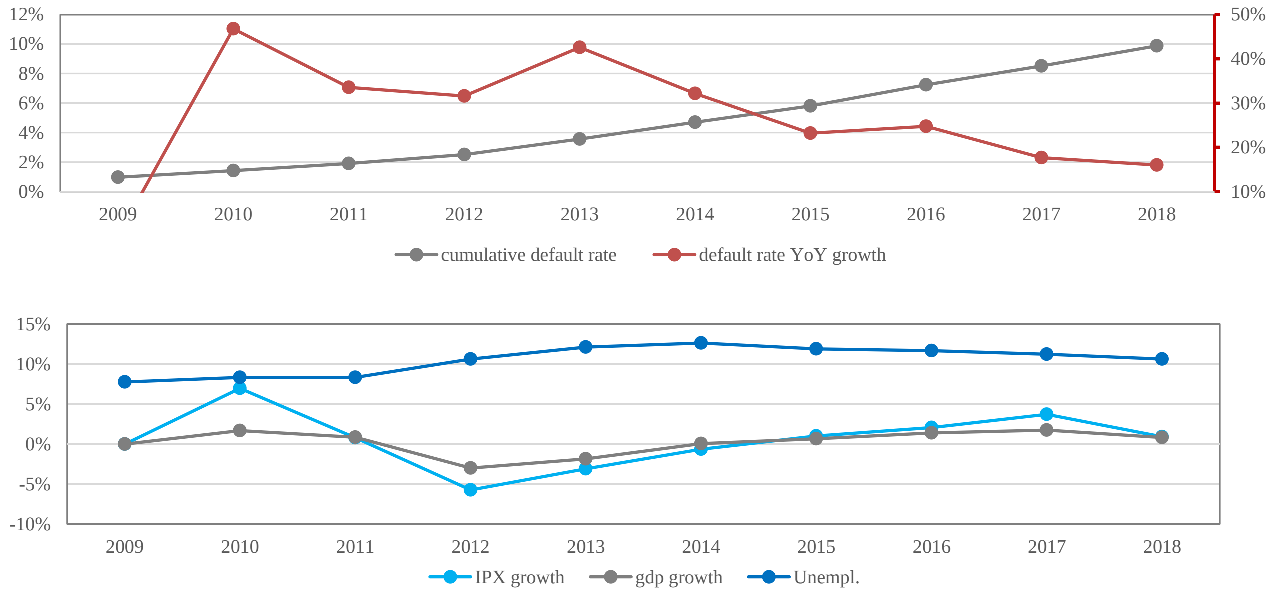
Before 2008, in the early stages of the crisis, Italy had been very marginally impacted, but the bankruptcy of Lehman Brothers in September 2008 shifted the focus from the sub-prime bonds and housing market, where Italian banks did not have material exposures, to the interbank loan market. The unexpected increase of interbank rates as well as the increased uncertainty of the financial environment and solidity of the counterparties, resulted in a reduced willingness to grant loans, and at the same time raising collateral requirements, to preserve liquidity. At this point the effects of the financial crisis reached the real economy. The pessimistic outlook induced the population to reduce spending, with a negative impact on real estate and consumer discretionary sectors. Businesses reacted to the reduced profit margins by cutting costs, via reduced full-time contract and not renewing temporary working arrangements.

Most affected parts of the economy had been younger and low-salary workers, and small and medium size enterprises. The former had minor guarantees protecting their job while SMEs were particularly impacted by the tendency of larger firms to internalize parts of the production processes that led to less business for sub-contractors, in most cases SMEs. At the same time, existing contracts were renegotiated to lower prices and payments to suppliers were delayed. The government response was to support banks and large companies to induce them to retain most of the workforce while cutting public spending, rather than increase taxes.

The origination of the 2011 crisis was instead driven by investors losing confidence in Italy's ability to repay its debt. As Italian government bonds yields increased above the 7% threshold, the servicing cost of debt was deemed to high compared to internal growth and investors started to close their exposures, driving a further increase of the yields. This escalation was partially mitigated by the ECB intervention as the Central Bank began to buy Italian bonds, succeeding in lowering the cost of debt. Moreover, given the persisting volatility in the markets and not so good forward-looking outlook, rating agencies also downgraded Italy sovereign debt by two notches to BBB+. The crisis de-escalated only after a change in the lead of the government and with austerity measures in place.

Afterwards the economy started its path to recovery, with GDP growth returning to positive, even if still low, with production levels and default rates normalizing – Table 12.

TABLE 12 In-sample default rate evolution and macroeconomic indicators in Italy



2.3 MODEL CALIBRATION

Three calibration techniques have been tested: logit regression on all the available variables, backward stepwise regression logit and neural network; results are presented in Table 13.

Accuracy results have been obtained with a repetitive cycle of training and testing of the model across the available historical data. The overall accuracy of ~70% for both type I and type II classification rates represents the average of model performances across the 2009–2018 time window, with different forecast horizons ranging from one to three years. In the second calibration cycle, macroeconomic data, more specifically the regional GDP per capita, GDP growth and GDP volatility were included as input in the regression. This affected marginally overall results in the logit and neural network model, and improved type II in-training classification accuracy in the stepwise-logit model. Out of sample testing had not been materially impacted; results are presented in Table 14.

Overall accuracy of the model is broadly aligned with similar research papers, with high number of firms and observations being used for calibration and where SME are well represented in the sample, as in Altman, Sabato and Wilson (Altman et al., 2010).

TABLE 13 Base Model – Accuracy

	Logit Model				**Logit and stepwise regression**				**Neural Network**			
	Training		Testing		Training		Testing		Training		Testing	
	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
Lag (t+1)	69%	73%	68%	73%	70%	72%	69%	73%	63%	49%	70%	42%
Lag (t+2)	69%	73%	67%	74%	70%	71%	67%	74%	70%	40%	66%	46%
Lag (t+3)	69%	72%	68%	72%	70%	71%	69%	71%	63%	46%	74%	39%
Average	69%	73%	68%	73%	70%	71%	68%	73%	65%	45%	70%	42%

TABLE 14 Model with Macroeconomic indicators – Accuracy

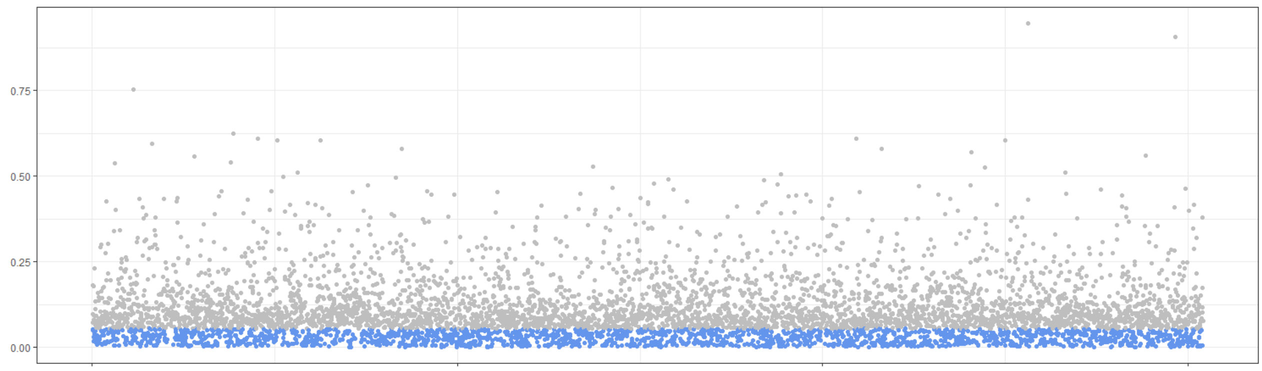
	Logit Model				**Logit and stepwise regression**				**Neural Network**			
	Training		Testing		Training		Testing		Training		Testing	
	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
Lag (t+1)	69%	73%	69%	73%	68%	74%	69%	73%	74%	37%	74%	37%
Lag (t+2)	69%	73%	69%	72%	68%	74%	70%	71%	60%	51%	68%	45%
Lag (t+3)	69%	73%	68%	71%	67%	74%	69%	71%	59%	53%	77%	37%
Average	69%	73%	69%	72%	68%	74%	69%	72%	65%	47%	73%	40%

Analysis of the results

Results are now analyzed at a more granular level, focusing on magnitude of the error and sub-buckets accuracies for defaulting firms. There is broad consensus in literature of asymmetric misclassification costs: the cost of starting a business relationship or lending a given amount of money with a client that might default on its debt is higher than not borrowing to a healthy firm. Therefore, we now focus our attention on the error rate of misclassification of defaulting firm:

- *Magnitude of the error*: the optimal cut-off point is calibrated with the aim of maximizing the accuracy for both sensitivity and specificity under the assumption that a certain amount of risk taking must be taken and that it would not be ideal for the lender to be either extremely conservative or to lend unconditionally. Error in the model's output is above a tolerance level of 5% for more than 90% of the incorrectly classified population and identifies all the situation where the accounting and financial ratios do not anticipate a situation of distress for the company. In Figure 4, the output from the model has been plotted and colored blue in case of estimation error.

FIGURE 4 Example of model output – 1 Year Lag, Defaulted population, cut-off point at 0.05 and type I accuracy at 70%



- *Bucket accuracy:* the model has been trained on a wide range of degrees of financial distress, as described in the previous section; error rates for each status bucket are presented in Table 15. The model performs relatively well for both early signs of distress ('*stato di insolvenza*') and more severe ones ('*fallimento*').

TABLE 15 Example of model output for buckets with more than 20 observations – 1 Year Lag, defaulted population, cut-off point at 0.05 and type I accuracy at 70%

Status details	True	False	Total	Error	Accuracy
Accordo di ristrutturazione dei debiti	50	67	117	57%	43%
Amministrazione straordinaria	90	20	110	18%	82%
Cancellata d'ufficio ai sensi art. 2490 c.c.	35	10	45	22%	78%
Cancellazione dal registro delle imprese	431	415	846	49%	51%
Cessazione delle attività nella provincia	10	15	25	60%	40%
Chiusura del fallimento	417	161	578	28%	72%
Chiusura della liquidazione	86	111	197	56%	44%
Concordato fallimentare	37	5	42	12%	88%
Concordato preventivo	137	106	243	44%	56%
Fallimento	2321	737	3058	24%	76%
Liquidazione	4	20	24	83%	17%
Liquidazione coatta amministrativa	44	6	50	12%	88%
Liquidazione volontaria	32	9	41	22%	78%
Provvedimento di cancellazione dal registro delle imprese	124	68	192	35%	65%
Scioglimento	20	22	42	52%	48%
Scioglimento e liquidazione	124	102	226	45%	55%
Stato di insolvenza	98	31	129	24%	76%
Trasferimento in altra provincia	25	16	41	39%	61%
Defaulted – Total	**4127**	**1955**	**6082**		

Financial ratios selection with Entropy Method

The entropy method is an alternative procedure to gauge indicators' predictive strength in the model. The underlying information theory states that lower degree of system's disorder can be linked to a small entropy value and, as the entropy increases, the associated explanatory power of the variable tends to reduce. This is used to estimate the discriminating ability for each attribute and select the more relevant ones for the purpose of bankruptcy prediction.

The database is a matrix containing *m* columns, corresponding to the financial ratios variables, and *n* rows containing the information for the *n-th* firm. Ratios are normalized to remove the impact from the average magnitude of the variables to calculate comprehensive entropy indicators when measurement units are not uniform.

$$(1) \quad x\_norm_{nm} = \frac{x_{nm} - x_{nm}^{min}}{x_{nm}^{max} - x_{nm}^{min}} \; with \; n = 1, 2, \dots N \; and \; m = 1,2, \dots, M$$

The proportion for each observation is calculated as:

$$(2) \quad p_{x(n,m)} = \frac{x_{nm}}{\sum_{n=1}^{N} xnm}$$

$$(3) \quad entropy_m = -k \sum_{m=1}^{M} p_{x(n,m)} * \ln{(p_{x(n,m)})}$$

Finally, the weight of each indicator is calculated as presented below, with d_j being information entropy redundancy.

$$(4) \quad w_x = \frac{d_m}{\sum_{m=1}^{M} d} \; with \; d_j = 1 - entropy_m \; and \; m = 1, 2, \dots, M$$

To test the appropriateness of the methodology, each one of the observations is ranked based on the distance from the ideal combination. This is set as the optimal value of every attribute from the weighted matrix calculated via entropy method. Results, presented in Table 16, show good differentiation in the average rank of the two classes of defaulted and non-defaulted, as the average position in the total sample for the former is higher than the latter.

Table 16 Ranking of observations based on TOPSIS technique (Technique for Order Preference by Similarity to Ideal Solution). Values range from 0.22 to 0.49, the smaller the closer to the ideal solution.

Class / Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Non-defaulted	0.3212	0.3214	0.3219	0.3215	0.3221	0.3224	0.3223	0.3229	0.3237	0.3243
Defaulted	0.3296	0.3278	0.3283	0.3265	0.3257	0.3245	0.3245	0.3255	0.3270	0.3320

2.4 CONCLUSIONS

The analysis performed on a sample of Italian firms active in the years from 2009 to 2018 needs to take into account the material presence of SME in the Italian landscape. The original assumptions of selecting indicators related to the three main accounting areas (financial, leverage and profitability) has been backed up by (i) the wide literature proposing similar solutions in default prediction modelling and (ii) by the model calibration performed on the sample used in this research. The results show a higher accuracy of conditional probability models, logit and stepwise logit, compared to neural network calibration. CPM models' accuracy is broadly aligned with the relevant literature. Lower levels of performance for the neural network calibration can be traced back to the heavy computational power required to run the calibration, as the routines are usually run on smaller samples, and hence the need to reduce the number of nodes during the cycles of training, which is likely to impact results. Satisfactory results of logit models constitute a solid basis on which the additional proposed enhancements, i.e. related to the inclusion of non-financial information and risk-taking propensities, can be implemented and more easily tested. This is mainly due to the upside of using CPM models allowing a clear understanding of the significance and contribution of each one of the selected variables to the default prediction.

3. DATASET ENHANCEMENTS

3.1 INTRODUCTION

In the chapter, three enhancements will be proposed to increase the *base model*'s performances and build a new integrated model which can leverage on additional data points of soft information to complement the evidence derived from accounting, financial and macroeconomic data.

The first enhancement is related to the addition of event data from online available information, the second is focused on analyzing the relationship with entrepreneurs' and firms' age, proposing a link with personal risk propensity's frameworks, and bankruptcy prediction. The latter can result specifically useful for industries where SME's presence represents a material portion of the population due to a more direct relationship between the entrepreneur's risk propensity and firm's performance. Finally, we will investigate the relationship between companies' default rates and the closeness to the local communities, where we have identified the central node in the local religious institutions.

3.2 EVENT DATA

In the previous section, the model has been calibrated with a wide range of accounting and financial indicators covering the different areas of each company balance sheet and P/L statement. In the second calibration cycle, macroeconomic data has been included with the aim of feeding information on the economic cycle and its impact on the default rates in the population. Both these explanatory variables, even though certainly related to the overall performance of the company, are lagging indicators, driven by the periodical, and not continuous, publication of the financial accounts and the macroeconomic data. This limitation can be overcome with the current dataset being enhanced with event data. Since all the available information from the AIDA database has been used, additional information can be sourced from public data and news available online.

For each company, the implemented model performs an online search with the legal name of the firm and consolidates all the information in text format to be analyzed. The addition of text analysis in default prediction modelling has also been proposed in other works (Stevenson et al., 2021). In this case, the author investigated the information available from the reports prepared by the loan officer in the initial credit assessment process, verifying whether the statements produced at origination can be predictive of default.

Generally, words' processing can be done following three main approaches: (1) thesaurus-based approach, (2) count-based approach and (3) inference-based approach calibrated with deep learning algorithms. For this research, a combination of (1) and (2) has been used, alternative (3) is more suited for direct impacts from news publication, for example stock prices going up or down. In the case of financial distress, the training of the algorithm would be slower as information published can be checked against annual information and not a continuous flow of prices. Therefore, the decision to code the model to analyze the text data against a list of sensitive keywords, generally used in news articles describing a situation of distress for the firm.

Code's description

TABLE 17 Web-searching function

```
news_search <- function(firm) {
  html_dat <- read_html(paste0("https://www.bing.com/search?q=",firm,"&FORM=HDRSC1&setmkt=it
  news_dat <- data.frame(
    Description =  html_dat %>%
      html_nodes('p') %>%
      html_text()
  )

  if (nrow(news_dat) == 0) {
    news_dat <- data.frame(Title = 'null', Description = 'null')
    news_dat$Title <-  paste(unlist(news_dat$Title), collapse = " " )
    news_dat$Description <-  paste(unlist(news_dat$Description), collapse = " " )
    news_dat <- news_dat[1,]
  } else {
    # Clean output and collapse in one row
    news_dat$Title <-  paste(unlist(news_dat$Title), collapse = " " )
    news_dat$Description <-  paste(unlist(news_dat$Description), collapse = " " )
    news_dat <- news_dat[1,]}
}
```

The code has been developed in R language and it is structured with a function to extract the information from the web pages and with the related script to elaborate the information. More precisely, the function's scope is to dynamically extract relevant search results for each one of the companies while the accompanying script elaborates the blocks of text provided by the function.

Target web pages are developed in HTML (hypertext markup language) which, as the name suggests, is not a programming language but rather a mark-up language used to build and structure the web page. Such pages are organized with tags, within “<>” brackets, that delimitate each section of the page. Each tag is accompanied by a similar one to indicate the beginning and the end of the section, for example a text paragraph would be identified by <p> at the beginning of the paragraph and by </p> at the end. Identifying the opening and closing tags is important to understand the structure of the web page as tags can also be nested in a tree-like structure and consequent hierarchy will be key to breakdown the page’s content.

The first input needed by the function is the link to the webpage containing the information we are looking for. For this purpose, there was the need of a dynamic link because we are looking iteratively for news data for different companies. This is achieved by examining how the link changes after each research. The first part is usually the main domain, in the case presented in Table 17 “search_engine.com/”. This is followed by the section we want to analyze. To increase the information being looked up by the function, two routines are run: the first one on the standard search results page, the second iteration limited to the news page only. The next part of the link contains the dynamic part of the vector, the company’s name, that is fed to the string. The final part is kept fixed as contains mostly formatting information.

After enabling the dynamic change of link by the function, the next step involves taking the relevant info from the webpage. This is achieved by extracting all the text paragraph in the search result page, which is, in this case, identified by the same tag, allowing a clean extraction just with one input in the function. The search algorithm already presents in this page the instances where the keywords we have specified in the research string appear in the result, therefore we can leverage this without going through every result’s link. The remaining part of the function has been built to clean the output format of the extraction to be more easily analyzed by the next block of code that searches and counts how many times the bankruptcy keywords defined in the previous section appear in the extracted block of text.

Results from enhanced dataset with event data

The new model has been trained and calibrated on a reduced sample of 50 firms, half of which defaulted to test classification accuracy of the keywords[2], and has then been deployed on the whole misclassified testing set of defaulted companies presented in Table 15. New implementation

[2] List of keywords coded in the model are: 'bankruptcy', 'bankrupt', 'default', 'bancarotta', 'liquidazione', 'concordato', 'crisi', 'fallimento'.

improved classification results, up to 5% for some of the sub-buckets, detailed results are presented in Table 18.

TABLE 18 Testing of the enhanced model

Status details	True	False	Total	Error	Accuracy	New Accuracy	%-change
Accordo di ristrutturazione dei debiti	50	67	117	57%	43%	43%	0%
Amministrazione straordinaria	90	20	110	18%	82%	84%	2%
Cancellata d'ufficio ai sensi art. 2490 c.c.	35	10	45	22%	78%	82%	4%
Cancellazione dal registro delle imprese	431	415	846	49%	51%	52%	1%
Cessazione delle attività nella provincia	10	15	25	60%	40%	40%	0%
Chiusura del fallimento	417	161	578	28%	72%	73%	1%
Chiusura della liquidazione	86	111	197	56%	44%	46%	2%
Concordato fallimentare	37	5	42	12%	88%	88%	0%
Concordato preventivo	137	106	243	44%	56%	60%	4%
Fallimento	2321	737	3058	24%	76%	77%	1%
Liquidazione	4	20	24	83%	17%	17%	0%
Liquidazione coatta amministrativa	44	6	50	12%	88%	88%	0%
Liquidazione volontaria	32	9	41	22%	78%	80%	2%
Provvedimento di cancellazione dal registro delle imprese	124	68	192	35%	65%	65%	0%
Scioglimento	20	22	42	52%	48%	50%	2%
Scioglimento e liquidazione	124	102	226	45%	55%	56%	1%
Stato di insolvenza	98	31	129	24%	76%	81%	5%
Trasferimento in altra provincia	25	16	41	39%	61%	66%	5%
Defaulted - Total	**4127**	**1955**	**6082**				

Overall, the model automatically reported 153 cases, out of the ~2000 firms, of text containing one or more mentions of the keywords. For what concerns the correct company identification and accurate negative news classification, out of the ~2000 firms, downloaded results pointing to the correct company in the testing database represents 73% of the total; negative news identification rules present 62% accuracy rate.

Comparison with findings from the systematic literature review

In chapter 1, a systematic literature review was performed to identify key features and downsides across the relevant literature. From this, the work from Altman, Sabato and Wilson (Altman et al., 2010) is specifically called out in this paragraph due to the similarities with the present study driven by the calibration technique being used, sample size and characteristics.

For comparability with the other studies the results have been presented aggregated and in different breakdowns. In Table 19, results are shown by region: the model returns broadly stable accuracy rates across north, center and south. The weighted average accuracy is slightly higher for the north driven by +2/4% in Type II accuracy compared to the other regions. This might be explained, albeit difference with the other regions is very small, by a lower overall default rate marginally impacting the calibration of the regressor coefficients during training. The underlying historical factors for this have been discussed in the first section of this work. Regarding the Type I error rate this is also broadly stable across regions, with accuracy rates also benefiting from the enhanced database with event data. The comparison with the systematic literature review's findings shows lower sensitivities (Type I accuracy) and specificities (Type II accuracy). The comparison with Altman, Sabato and Wilson (Altman et al., 2010) shows though a slightly higher specificity compared to the published results.

To better understand the differences in type and type II error rates, in Table 20, results are instead presented by Tier, ranging from 1 to 3 depending on the severity of the company financial distress. One key difference between the presented results and most of the other papers is the default definition. While this is usually overlapping with the judicial definition of bankruptcy, in this research the scope is wider as it includes also earlier stages of financial distress. Under Tier 1 only the most severe stages are included: (i) *amministrazione straordinaria*, (ii) *concordato fallimentare*, (iii) *fallimento*, (iv) *liquidazione* and (v) *liquidazione coatta amministrativa*. In Tier 2, company's classifications being included are (i) *accordo di ristrutturazione dei debiti*, (ii) *chiusura del fallimento*, (iii) *chiusura della liquidazione*, (iv) *concordato preventivo*, (v) *liquidazione volontaria*, (vi) *scioglimento e liquidazione* and (vi) *stato di insolvenza*. These are less severe stages compared to Tier I, but all are indeed flagging a situation of liquidity shortage, but with a slightly higher recovery probability. In Tier 3 all the remaining stages are bucketed, the full list of in-sample classification has been presented earlier in the thesis.

The model can better classify more consolidated stages of distress, mostly aligned with the international definition of bankruptcy. Reasons can be mainly driven by the usage of financial ratios as the main input in the model. These are, by definition, lagging indicators; therefore, a very early situation of reduced liquidity buffer can be harder to identify. In Table 20, it is shown how comparable results from the quantitative literature review are aligned to the ones from our model. The results exceed expectations compared to the relevant literature analyzed in the first chapter because of (i) mixed and (ii) nonmatching random training sample, (iii) high heterogeneity in the Italian enterprises landscape. Among the main objectives, the model should be applicable to all firms, and this is indeed the case as no carve-outs have been done to the training sample: both listed and SME are included and results prove that both sensitivity and specificity wise the model performs in line with published results. Secondly, the sample corresponds to a random sample extracted from all the available information in the market which can then be assumed to correctly represent the actual population. It should be expected that smaller samples would be easier to manage to delimitate the scope and better define the specific targets, restricting the research to a specific industry or firm's type. It is believed that the chosen sample is more in line with modern needs of classification and initial due diligence that require higher flexibility of scope and that can also flag early situation of distress that can impact future projection of recovery on granted loans. Finally, on in-sample heterogeneity, this is the case for the Italian landscape, the model performs accurately across buckets, regions and firm's structures. The model is specifically suited where an accurate and relative quick and automated analysis need to be performed, where the information on the distance from safe benchmarks can be easily calculated; for example, when a bank is deciding whether or not to grant smaller to mid-size loans, which is often the case in the Italian market. Moreover, extracting information from the public sources online, as described in the previous section, can bridge the gap derived from accounting and financial ratios and their inherent lag. This can, of course, drive better decisions upfront, i.e. on whether or not to grant a loan, but can also accelerate the recovery process by providing an easy to use and automatic tool in the monitoring phase of the credit cycle.

Final consideration on the calibration methodology. Results show lower accuracy under NN compared to the other conditional probability models (CPM) alternatives. One driver may be that NN implementation generally performs better with smaller samples than the one used in this research. Despite a slightly lower accuracy discovered in the mentioned quantitative literature review, we believe that the CPM implementation can be a more useful tool because of the explanatory power and easier output interpretation of the technique itself compared to the black-box structure of neural networks. This allows to breakdown more granularly where the weaknesses of the firm are compared to the sample's benchmarks, allowing a closer specific monitoring or fine-tuned solutions, i.e. longer loan maturity or smaller overall ticket, to account for these.

TABLE 19 Results comparison with systematic literature review

				Literature review		Altman, Sabato, Wilson 2010	
Model	**Type I**	**Type I (Enhanced database with event data*)**	**Type II**	**Type I**	**Type II**	**Type I**	**Type II**
L	**68%**	**70%**	**73%**	**77%**	**81%**	**77%**	**73%**
North	68%	69%	74%				
Center	65%	67%	72%				
South	68%	70%	70%				
L+M	**69%**	**71%**	**73%**				
SWL	**69%**	**70%**	**73%**	**77%**	**81%**	**77%**	**73%**
North	68%	69%	74%				
Center	65%	67%	72%				
South	69%	70%	69%				
SWL+M	**69%**	**71%**	**73%**				
NN	**70%**	**71%**	**42%**				
NN+M	**74%**	**76%**	**37%**	**88%**	**84%**		
e	**68%**	**70%**	**66%**	**77%**	**81%**		

**Estimated improvement based on iteration on 6082 observations*
(L = Logit, SWL = Stepwise Logit Regression, NN = Neural Network, M = addition of macroeconomic data, e = entropy method to select relevant indicators)

TABLE 20 Results comparison with literature review – Type I error rate by sub-bucketing

			Literature review		Altman, Sabato, Wilson 2010	
Model	**Type I**	**Type I (Enhanced database with event data)**	**Type I**	**Type II**	**Type I**	**Type II**
L	**68%**	**70%**	**77%**	**81%**	**77%**	**73%**
Tier 1	76%	77%				
Tier 2	62%	63%				
Tier 3	54%	55%				

3.3 ENTERPRENEUR'S AGE

As with some other research fields, the risk associated with a particular firm is also dependent on a series of qualitative factors strongly linked with personal propensities and attitudes. In the field of corporates' bankruptcy prediction, a direct relationship between the actions of each individual and the influence on either the final output or the performance of the firm can be sometimes hard to identify. Especially in large corporations, where different agents are operating at the same time, it can be difficult to identify the driving forces, breaking them down to the individual level. Moreover, a structural oversight, both vertically, i.e. the internal collaboration within the teams with directors and employees, and horizontally across different departments, for example Risk or Compliance, can have a mitigating effect on personal drivers and behavioral biases.

Due to the mentioned factors, small enterprises represent a more feasible target of the study as it can be assumed that, given the smaller management group, smaller enterprises are in most cases family owned, the performance and amount of risk taking have a more direct link to the individual propensities of the entrepreneur. The sample used in this section has been restricted to ~4800 firms' observations from Italy presenting a physical person fiscal code ("*codice fiscale")* from where the entrepreneur's attributes have been estimated.

The first factor included in the analysis is the age of the entrepreneur. Across the literature, the definition for risk propensity can vary and can also depend on the local culture and, more widely, socially accepted behaviors. For example, in the studies of young generations and adolescents, risk taking can be defined as engaging in actions or activities that might compromise the wellbeing, or health, of another individual, or group of people. A downside of this definition, in this case causing specific harm, is that it cannot be easily extended across societies and cultures. For example, conducting a Western style of living can be perceived as normal in some countries and considered risky in other Regions of the world (Duell et al., 2018). Moreover, this definition can be hardly linked to firm's performance and considered accurate predictor of financial distress. Therefore, the characterization that will be used in this research is wider and defined as the inherent propensity of risk taking or the personal inclination to take risks.

The underlying personality traits are usually connected with biological factors and influenced from the surrounding environment: risk taking serves as a fitting role in the society (Mata et al., 2016). Among the main theoretical structures where this can be framed, there is the life-history theory: individuals are inherently driven to maximize their fitness in the society and allocate their time and energy for this purpose. This is strongly linked with the sensation seeking framework, that focuses on the analysis of personal needs of feeling of excitement or novelty, variety seeking or any other stimulus to maintain an optimal level of activation. More precisely, this has been defined (Zuckerman, 2007) as a “trait defined by the seeking of varied, novel, complex, and intense sensations and experiences, and the willingness to take physical, social, legal, and financial risks for the sake of such experience”.

The two variables chosen within the scope of this analysis are age and surrounding environment. Starting with the latter, this traces back to the life-history theory. A worse state of economy and macroeconomic conditions can lead individuals to take more risks at an earlier stage of life compared to other societies. On one side, life expectancy needs to be accounted for, on the other side the need to adapt, or fit in, can be accelerated by a harsher environment. At the same time this can be offset by life-events happening at an earlier stage. For example, in some societies starting parenting or employment comes on average at an early stage of life and may leave an imprinting of personality maturation that can reduce innate risk propensity and skew the comparison simply by age group across regions.

On the age of the individual, there is ample evidence in literature of an inverted U-shaped pattern with greater risk taking in the first part of life, reducing as the individual grows older (Duell et al., 2018). The shape of the curve raises some questions on the complete adaptability of the life-theory framework, in particular the monotonic declining trend after the peak. It should be expected that older individuals, with a higher mortality rate compared to the younger generations, would be willing to take on more risks in the hope of more immediate achievements. In practice, this is not always the case as studies have confirmed lower levels of risk taking across older individuals, within the life-history theory this can be justified by the possibility of transfer of the life-built resources to the close family or future generations.

The status of the firm, defaulted or not, and the total revenues are taken into account in order to test the applicability of the framework on the selected group of micro, small and medium enterprises. From a higher risk propensity of the company higher returns can be expected, but also a higher default rate. This is indeed the case in the sample, as shown in Figure 5. The default rate is inversely correlated with the age variable, reaching the peak for less than 30 years old category, and then following a declining trend. Moreover, looking at the distribution of revenues across the performing companies of the sample, it is also found a high upside potential in the first bucket, entrepreneurs aged 30 or less, which is broadly aligned to those of the older generations. These ones though, have a higher in-sample average. This is in line with expectations as risk taking is usually correlated with lower performances, in this case lower revenues, in smaller enterprises. This is because of assumptions of less structured internal and external monitoring and risk management of risk-taking strategies. These are usually based on the intuition of a small number of people, sometimes the founding family or entrepreneur, without deep systemic analysis or external feedback as well as less external pressure, i.e. auditors or other stakeholders, to motivate or justify the chosen strategy.

FIGURE 5 Default distribution by age of the entrepreneur

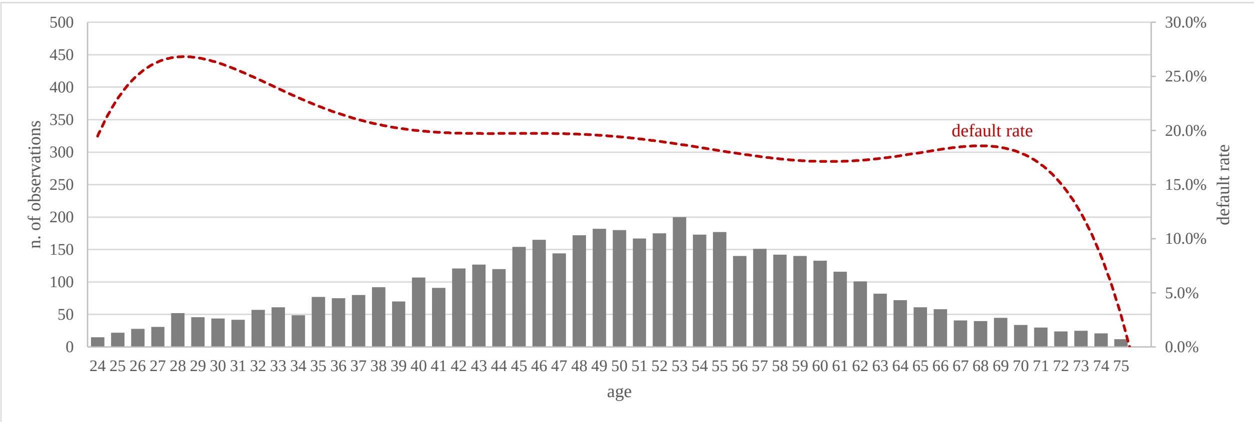
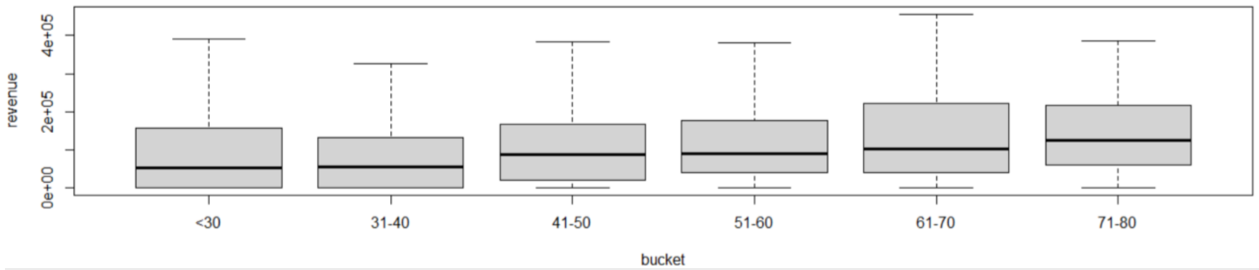


FIGURE 6 Revenues distribution by age bucket



3.4 FIRM'S AGE

The risk-taking propensity of the entrepreneur can also be studied from the angle of the propensity to innovate and proactiveness (De Massis et al., 2014). As mentioned before, the sample of analysis includes only micro, small and medium enterprises. This is an important point to stress since the first distinction found in the relevant literature is between family-owned companies and non-family owned. Family firms are less inclined to take risks compared to the other types of companies. There are various factors contributing to this tendency. Among those, agency theory demonstrates an inverse correlation between equity ownership and risk-taking propensity. Secondly, since most of the family wealth is usually invested in the company, strategic or investment decisions can be more difficult to go through. Moreover, additional influence comes from concerns on the ability to preserve the social wellbeing of the closer future generation as well as the good reputation of the family name.

There are many variables identified with a higher risk tendency, but all of them are correlated to the two main categories of innovativeness and proactiveness. The first is mainly related to output innovation, i.e. releasing new products or services, the latter as being pioneers in the industry where the firm operates. Both are key factors of one company's longevity. If the investment is successful, the enterprises can gain a big advantage over competitors and can also have a driving force in shaping the external environment. When management is reduced to a small number of people, with family involvement or friendships among founding members, literature is divided on whether proactiveness and innovativeness are encouraged or not. This is because of the alternative assumptions leading to agency or stewardship behaviors.

On one side, the agency theory assumes that, especially for larger families, business goals are not aligned to family objectives. This means that management is driven by self-interest and the main force would be to preserve the individual's wealth rather than to invest on the firm longevity and future. Consequently, expectation under this theory would be that firms focus on inflated management compensations at the expense of research and pursuing of new opportunities. On the other side, the stewardship theory proposes converging alignment between family and company goals. Family relationships are usually driven by altruism, affection and trust and this is mirrored to the commitment to the wellbeing of the enterprises. This means that the company represents an extension of the family, long-term survival and reputation are key priorities because those become attributes and characterizations of the family itself, therefore innovativeness and proactiveness represent a strong commitment under this assumption. Given the complexity of human nature the two theories can be both valid at different stages of the company's life. In Figure 7, the trend on this relationship is presented (De Massis et al., 2014).

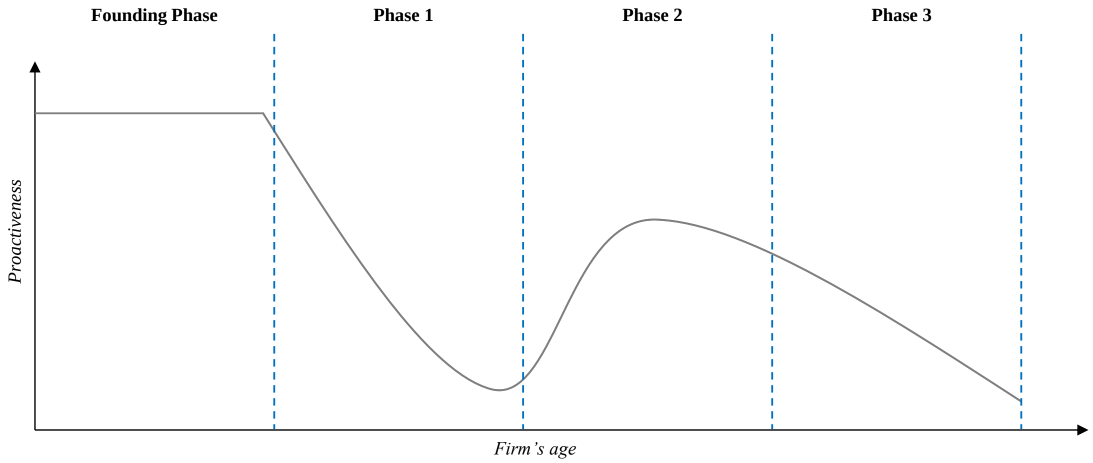
This framework describes the propensity as driven by the life cycle of the enterprises. During the founding phase there is high alignment of goals and collaboration within management, everyone has high ambitions and expectations for the newly created company. Altruism and informality are key factors of this first phase: there is good faith in everybody's intentions of having the well-being and development of the company as key priority, aiming to minimize everything that may subtract resources to the firm.

In the next phase, the firm has now successfully passed the initial stage. This consolidates the positive feelings mentioned among founding members adding on top also feelings of pride and prestige due to having established a working business. As the number of members that rely on the company grows though, management goals can get increasingly fragmented and each member's goal can diverge from what would be in the best interest of the company as a whole. More formal agreements start to be needed as trust among members starts to decline. This has a negative impact on firm's ability to innovate as management gradually turns from stewards to agents more eager to further pursue conservative policies and strategies with special attention to efficiency rather than innovation.

The firm loses that competitive advantage that characterized its founding phase while the goals of each member of management begins to diverge from the optimal strategy of the company, leading to reduced performance overall. This is the driving force of the following realignment among each individual's goal to what would be best for the company. As the monetary contribution distributed reduces, a new shifting focus towards innovation and proactiveness can find more consensus. Objectives begin to converge again with the reemerge of altruism and trust as the agent behavior is abandoned to returns to a stewardship attitude among founding members.

After this phase, the firm has now been operating for a long time and it is also reasonable to assume to have a well-established position in the competitive landscape. The trends at this stage are usually to open management to external components, with stronger internal controls and monitoring. As new standard corporate's routines and policies take increasing weight in strategic decisions, family ability to drive and lead change within the company becomes weaker. In this stage, founding members main concern is ownership and a more conservative behavior, with another shift from stewardship to agent propensity, arises again reducing the innovativeness push, which is now on a declining trend inversely related to firm's age.

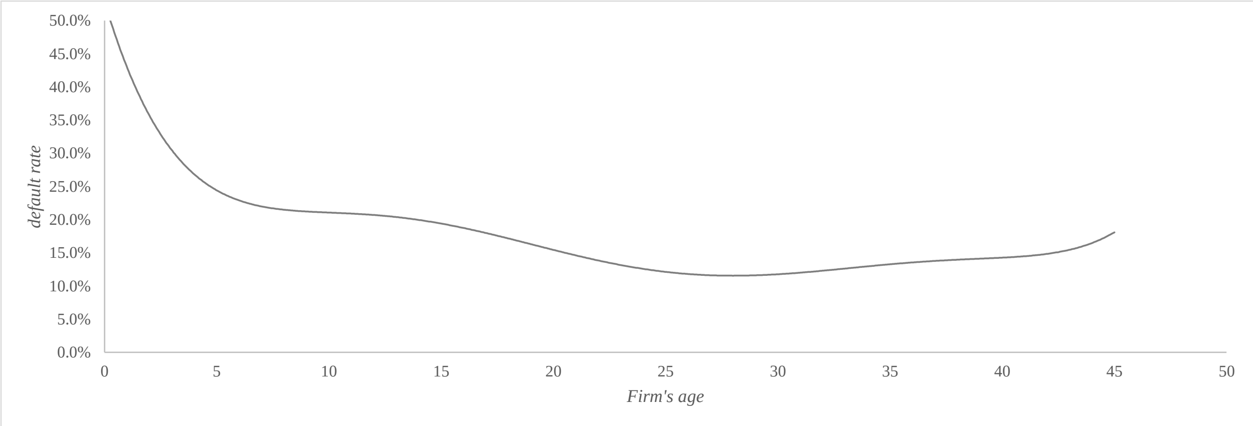
FIGURE 7 Proactiveness trend (A. De Massis et al – 2014)



Proactiveness	High	Declining	Growing	Declining
Expected behavior	Stewardship	Agency	Stewardship	Agency
Dynamics	▪ Goals overlapping and convergence ▪ High trust and altruism among management	▪ Goals start to diverge ▪ Reduced trust and increasing need of formality	▪ Goals begin to re-align to offset reduced performance	▪ Declining push towards proactiveness due to management focus mainly on ownership

Testing the described framework on the selected sample of enterprises, we found consistency with the mentioned attributes as presented in Figure 8. This is also consistent with what was presented in the previous section on entrepreneur's age as would be expected considering that newly born firms will belong to younger founders.

FIGURE 8 Default rate distribution by firm's age



The assumptions presented above and tested on a reduced sample are now applied to the wider population referenced in the previous chapters. Given the aligned trend for the entrepreneur and firm age and due to information availability constraints, only the latter has been considered in the following simulations. As mentioned, the correlation among the variables can be explained because younger firms are usually started by younger entrepreneurs and vice-versa for older firm's and their leadership. Improvements are mainly impacting model's specificity as can be seen in Table 21.

TABLE 21 Model accuracy for micro and small enterprises with less than 50 employees and revenues below €10M, with and without age variable

	Stepwise Logit Model (with age variable)				**Stepwise Logit Model (w/o age variable)**			
	Training		Testing		Training		Testing	
	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
Lag (t+1)	64%	69%	60%	73%	63%	69%	60%	71%
Lag (t+2)	65%	69%	61%	69%	64%	69%	61%	68%
Lag (t+3)	65%	69%	59%	69%	64%	69%	59%	68%
Average	65%	69%	60%	70%	64%	69%	60%	69%

As a second step, to further test the applicability of the framework, another analysis is performed in order to study the significance of the age value while changing the firm's dimension. This is achieved by running the model on different partitions on the sample obtained by changing the number of people employed and the total revenues. The first one above the EU definition of small and medium enterprises (SME), with a threshold of more than 250 employees and revenues above €50M, the second one below the mentioned thresholds. The age variable is significant only below the SME threshold, in line with the theoretical assumptions.

TABLE 22 Summary statistics – logit regression

Less than 250 employees and below €50M of revenues						More than 250 employees and above €50M of revenues					
	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**			**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**	
(Intercept)	14.260	1.17E+02	0.122	0.90297		(Intercept)	-3.858	4.71E-01	-8.2	2.41E-16	***
liquidity ratio	0.081	3.51E-02	2.317	0.02049	*	liquidity ratio	-0.203	1.47E-01	-1.387	0.165433	
short-term debt ratio	-15.890	1.17E+02	-0.136	0.89192		short-term debt ratio	1.592	4.33E-01	3.677	0.000236	***
debt ratio	0.000	9.05E-05	0.976	0.32899		debt ratio	-0.002	8.60E-04	-1.903	0.056998	^
fixed-asset (financial) coverage ratio	0.000	4.15E-04	0.961	0.33639		fixed-asset (financial) coverage ratio	0.003	1.20E-03	2.093	0.036381	*
bank debt/revenues ratio	0.002	1.29E-03	1.467	0.14228		bank debt/revenues ratio	0.016	3.61E-03	4.413	1.02E-05	***
cost of credit	0.051	5.40E-03	9.426	< 2e-16	***	cost of credit	0.094	1.44E-02	6.515	7.28E-11	***
interest coverage ratio	-0.001	5.63E-04	-0.976	0.32897		interest coverage ratio	-0.007	3.08E-03	-2.226	0.026002	*
interest/revenues ratio	0.021	4.18E-03	5.134	2.83E-07	***	interest/revenues ratio	0.132	2.30E-02	5.727	1.02E-08	***
equity ratio	-0.032	1.69E-03	-18.683	< 2e-16	***	equity ratio	-0.033	5.23E-03	-6.394	1.61E-10	***
equity ratio (from external stakeholders)	0.030	5.33E-03	5.66	1.51E-08	***	equity ratio (from external stakeholders)	0.025	3.70E-02	0.679	0.497357	
net financial position	0.000	1.67E-09	4.421	9.81E-06	***	net financial position	0.000	9.94E-10	-2.372	0.017699	*
debt/equity ratio	-0.001	4.80E-04	-1.91	0.05614	^	debt/equity ratio	0.002	4.69E-03	0.486	0.62669	
debt/EBITDA ratio	0.000	6.27E-04	-0.131	0.89564		debt/EBITDA ratio	0.002	1.76E-03	0.9	0.367862	
EBITDA	0.000	1.31E-05	-4.972	6.64E-07	***	EBITDA	0.000	3.16E-06	-2.37	0.017783	*
EBITDA/Sales ratio	-0.003	8.69E-04	-3.275	0.00106	**	EBITDA/Sales ratio	0.001	3.67E-03	0.149	0.881274	
ROA	0.006	2.79E-03	1.975	0.04822	*	ROA	0.036	1.17E-02	3.05	0.002291	**
ROI	-0.027	2.88E-03	-9.436	< 2e-16	***	ROI	-0.026	8.64E-03	-3.025	0.002489	**
ROS	-0.002	4.08E-03	-0.395	0.6927		ROS	-0.036	1.35E-02	-2.679	0.007392	**
ROE	0.000	8.72E-04	-0.23	0.81841		ROE	-0.002	2.51E-03	-0.993	0.320896	
non-core losses/gains	-0.003	3.89E-04	-8.313	< 2e-16	***	non-core losses/gains	-0.002	8.28E-04	-2.537	0.011168	*
age	-0.003	1.54E-03	1.652	0.09858	^	age	0.004	3.45E-03	1.274	0.202709	

$^\wedge\ p < 0.10$; $*\ p < 0.05$; $**\ p < 0.01$; $***\ p < 0{,}001$

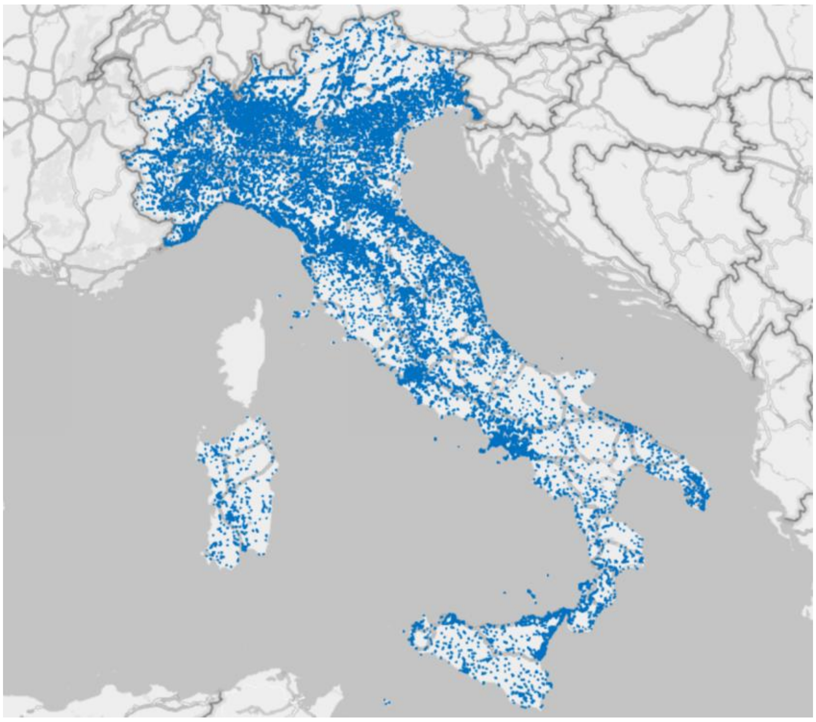
3.5 CLOSENESS TO LOCAL COMMUNITIES

We propose a framework to leverage and explore the links between the companies' performance and the closeness to the local communities. For this purpose, religious institutions have been identified as the central point of societies and we investigate whether the distance from this central node can have any effect on performance and default. As for the previous section, the link and the impacts on economic and financial variables is assumed to be better identifiable for small and medium enterprises due to stronger economic ties with local communities than in larger enterprises. Many papers have already investigated the influence of the local religious presence in affecting institutional and political change (Belloc et al., 2016), economic growth (Barro and McCleary, 2003) and enhancing social networks and group cohesion (Wald et al., 1990). Building on the latter, we propose a framework that leverages on the distribution of local churches to estimate the positive spillover effect of the closeness to the center of the community, as mentioned above, identified in the local church. In Italy, the proposed framework also benefits from the limited religious plurality due to historical factors (Barro and McCleary, 2003), hence making acceptable the use of only catholic reference points. Therefore, we proceed to test our hypothesis that the closeness of the enterprises to the local community, identified in the closest religious institution, impacts company's performance and the default rate.

We have compiled a local database with the list of all the available institutions – with this term we refer to the Italian terms *abbazia, diocesi* and *arcidiocesi*[3] – with public available information as of the date of this study. As presented in the figure below, the distribution is capillary across the country.

[3] Data has been extracted from https://www.chiesacattolica.it/annuario-cei/ricerca-parrocchie/

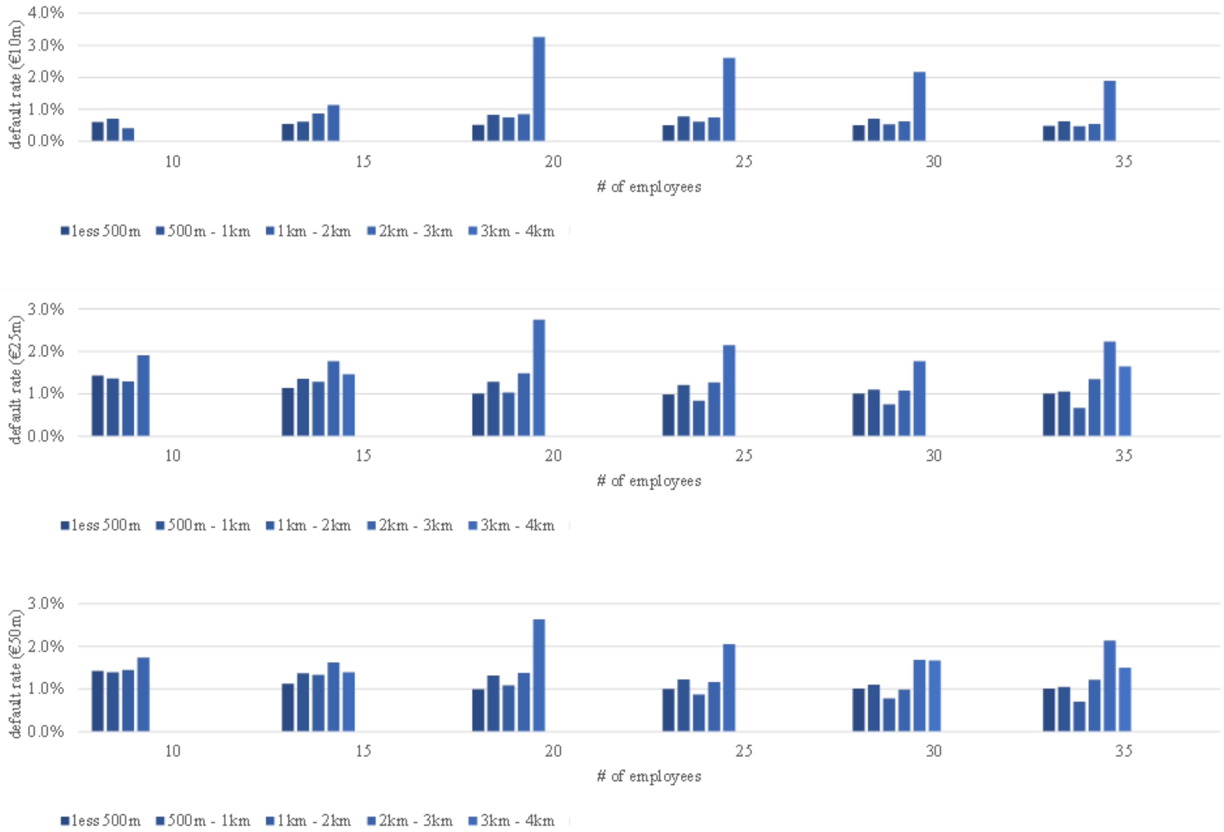
FIGURE 9 Distribution of religious institutions in Italy (Sample size: ~17,000)



From a computational point of view, the religious institutions and the companies addresses have been converted to the correspondent latitude and longitude coordinates. Then, we have matched each enterprise with the closest church, being the point of minimum distance between the enterprises and any religious location in the database. Distance has been computed with the *geosphere* R-package, methodology assumes a spherical earth, not taking into account any ellipsoidal effects. A sensitivity analysis of the default rate of the Tier 1 bucket[4] has been performed by comparing the default rates with the estimated distance to the local church. Three analyses are presented with revenue levels of €10 million, €25 million and € 50 million, and number of company's employees ranging from 10 to 35. Due to sample constraints, results are presented only if the available number of enterprises for the identified cluster is greater than 1000. Results, presented below, appear to support our assumption that, for firms with revenues below €50 million and with number of employees below 35, to a higher distance from the local religious institution corresponds an average higher default rate.

[4] As described in the previous chapter of the study, the tier 1 bucket includes only the most severe stages of defaults: *(i) amministrazione straordinaria, (ii) concordato fallimentare, (iii) fallimento, (iv) liquidazione and (v) liquidazione coatta amministrativa*

FIGURE 10 Sensitivity analysis of Tier 1 default rate distribution across difference distances, on number of employees (x-axis) and revenues (y-axis); results have been presented only if available sample is greater than 1000 companies



We then proceed to feed the new variable to the classification model described earlier to check for significance and enhancements to the model's performance. Model's accuracy with the inclusion of the distance variable is higher on average, considering both type I and type II error rate, hence presenting more balanced results. The results presented below represent the average accuracy of the model throughout the running cycle.

TABLE 23 Model accuracy for micro and small enterprises with less than 50 employees and revenues below €10M, with and without distance variable

	Stepwise Logit Model (with distance variable)				**Stepwise Logit Model (w/o distance variable)**			
	Training		Testing		Training		Testing	
	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.	Sens.	Spec.
Lag (t+1)	69%	65%	64%	64%	67%	63%	70%	56%
Lag (t+2)	69%	66%	66%	57%	68%	63%	63%	60%
Lag (t+3)	69%	66%	68%	52%	68%	63%	73%	45%
Average	69%	66%	66%	58%	68%	63%	69%	54%

TABLE 24 Summary statistics – logit regression

	Estimate	Std. Error	z value	Pr(>|z|)	Significance
(Intercept)	-3.940e+00	9.802e-01	-4.020	5.82e-05	***
liquidity_ratio	-1.519e-01	3.579e-02	-4.243	2.20e-05	***
debt_ratio_short	2.465e-01	9.806e-01	0.251	0.801509	
debt_ratio_long	5.413e-01	9.880e-01	0.548	0.583809	
depreciation	-1.262e-01	1.943e-02	-6.495	8.29e-11	***
debt_ratio	1.824e-04	1.204e-04	1.514	0.129904	
coverage_assets_financials	-1.208e-03	6.661e-04	-1.814	0.069676	.
bank_debt_revenues_ratio	-1.757e-02	1.802e-03	-9.749	< 2e-16	***
cost_of_credit	-2.661e-02	7.199e-03	-3.696	0.000219	***
negative_interest_coverage	-1.853e-03	5.550e-04	-3.339	0.000840	***
borrowing_costs_revenues	-2.056e-02	6.671e-03	-3.082	0.002055	**
financial_independence	1.758e-03	1.481e-03	1.187	0.235120	
degree_ind_third_parties	8.888e-03	6.071e-03	1.464	0.143186	
net_cash_position	1.131e-08	1.712e-09	6.606	3.96e-11	***
debt_equity_ratio	1.189e-04	7.665e-04	0.155	0.876779	
debt_EBITDA_ratio	-2.493e-03	5.743e-04	-4.341	1.42e-05	***
EBITDA	-9.420e-05	2.353e-05	-4.004	6.23e-05	***
EBITDA_sales	-2.890e-03	6.200e-04	-4.661	3.14e-06	***
ROA	-6.564e-03	3.234e-03	-2.030	0.042403	*
ROI	-3.051e-02	3.164e-03	-9.642	< 2e-16	***
ROS	8.566e-03	4.705e-03	1.821	0.068644	.
ROE	4.806e-03	1.154e-03	4.166	3.09e-05	***
interest_expense_income_ratio	-2.338e-03	3.850e-04	-6.071	1.27e-09	***
distance	1.884e-04	2.295e-05	8.207	2.27e-16	***

^ $p < 0.10$; $* \; p < 0.05$; $** \; p < 0.01$; $*** \; p < 0{,}001$

3.6 CONCLUSIONS

The attempts to further enhance the predictive ability of the model by enriching the dataset have been successful in improving the overall accuracy of the classification. The bridge with theoretical frameworks mostly belonging within the behavioral finance stream of research can provide new points of views other than the consolidated ones. In addition to this, we have leveraged existing contribution linking economic development to the closeness to the social centers, in our case local churches. Our analysis has also confirmed a degree of statistical significance to this factor, providing additional variables that can be fed to the model.

Finally, we present below the breakdown of the results between the base and the enhanced models, noting that: (i) the differences with the tables of comparison presented in chapter 2 are due to the sample of analysis being a carve-out of micro and small enterprises and (ii) the analysis presented takes into account the testing on the last available year in the sample compared to the average of the model's run presented in Table 23. As presented, the accuracy gains reported by the model running on the enhanced dataset are improving the overall performance. This is achieved with a higher sensitivity in identifying defaulted firms that more than offset the losses in specificity.

TABLE 25 Type I and type II accuracy rates for the Base Model and the Integrated Model (Lag 1, hold-out sample for testing corresponding to 2018)

	Base Model		Integrated Model			Difference	
Model	**Type I**	**Type II**	**Type I**	**Type I (Enhanced database with event data*)**	**Type II**	**Delta Type I**	**Delta Type II**
SWL	**49%**	**77%**	**60%**	**61%**	**70%**	**13%**	**-7%**
North	50%	78%	61%	63%	71%	13%	-7%
Center	47%	74%	56%	58%	66%	11%	-8%
South	47%	73%	60%	61%	66%	14%	-7%

	Base Model		Integrated Model			Difference	
Model	**Type I**	**Type II**	**Type I**	**Type I (Enhanced database with event data*)**	**Type II**	**Delta Type I**	**Delta Type II**
SWL	**49%**	**77%**	**60%**	**61%**	**70%**	**13%**	**-7%**
Tier 1	49%		55%	56%		8%	
Tier 2	49%		68%	69%		20%	
Tier 3	48%		61%	63%		15%	

**Estimated improvement based on iteration on 6082 observations*

4. CONCLUSIONS

The objective of this study was to propose an integrated model for default prediction whose results have demonstrated to improve the overall accuracy rates of the base line model, which is mainly leveraging on accounting, financial and macroeconomic indicators.

For the base-line exercise, three models have been calibrated, leveraging on the most common techniques – logit, stepwise logit and neural network – with accuracy rates being overall aligned to the relevant literature, whose relevant references have been selected through the systematic literature review. The models have been tested on a wide range of definitions of default, ranging from the most severe ones to lighter degrees of financial distress; as expected, more advanced stages of distress are more easily identified. The next steps have been focused on enhancing the models' performances, the objective has been tackled from two different angles: the enhancement of the dataset with event data, and by analyzing behavioral dynamics of companies and entrepreneurs which specifically impacts micro, and small enterprises, constituting a material portion of the research sample.

The first layer to build the new integrated model is based on elaborating the online publicly available information. To extract this data, a routine has been developed on news search engine and news websites. From these, the information is downloaded as text which is then analyzed word-by-word against a vocabulary of relevant keywords. The implementation impacted positively across all the sub-buckets of default, with an increase in Type 1 accuracy up to +5% in the sample. Then, the attention has been focused on testing the behavioral propensities of entrepreneurs and how this can affect their risk-taking attitude and, consequently, the enterprise's performance. The sample for this analysis has been reduced to only include SMEs due to the assumption of stronger relationship between personal risk propensities and impact on performances.

The frameworks related to the entrepreneurs age and firm's propensity to innovate – adapted as a point in time function of the company's lifecycle – were deemed relevant to the Italian sample through backtesting literature assumptions on the in-sample risk-return distribution. Through this implementation most of the improvements in accuracy were related to type II error rate in our testing. Finally, we have proceeded to investigate the links between the local communities and SME on similar assumptions described in the previous paragraph. The center of the local community has been identified in the closest religious institutions and the distance from this point has been calculated for all the companies selected in the sample. Results with the described distance variable, i.e. from the closest church, provide a better overall accuracy of the model.

The integrated model, running with all the described additional three layers, results in higher accuracy in type I, improving base line performance by 13% on average across the different default stages, more than offsetting the rebalancing in overall accuracy due to the reduced specificity.

The improvements in predictions accuracy have been mainly driven by the inclusion of additional layers of soft information in the model. We believe that future research efforts should aim to further explore potential benefits from qualitative data, especially when samples include a material portion of SME. For example, this can be achieved by extracting information not only from online news portals, but also from social media websites. Particularly when looking at personal risk propensities, likes to certain groups or posts can be directly related to higher risk tendencies with implicitly higher default rate. In addition to this, further research angles may also explore the potential presence of systemic risks across the companies in the sample. In this case, the default of one company can be a warning sign for the going concern of similar companies in the sample. This can be particularly relevant when assessing the likelihood of default of networks of companies exposed to similar risks.

5. ANNEX – REFERENCES, LIST OF FIGURES AND TABLES

REFERENCES

Agarwal, P., Patni, I., 2019. Applicability of Altman Z-Score in bankruptcy prediction of BSE PSUs. Journal of Commerce and Accounting Research 8, 93–103.

Alaka, H., Lukumon, O., Hakeem, A., Kumar, V., Saheed, O., Olugbenga, O., Bilal, M., 2017. Systematic review of bankruptcy prediction models: towards a framework for tool selection. Expert Systems with Applications 94, 164–184.

Allen, W., Moessner, R., 2013. The liquidity consequences of the euro area sovereign debt crisis. BIS working papers 390.

Almamy, J., Aston, J., Ngwa, L., 2016. An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: evidence from the UK. Journal of Corporate Finance 36, 278–285.

Almeida, L., 2023. Risk and bankruptcy research: mapping the state of the art. Journal of Risk and Financial Management 16, 1–18.

Altman, E., 2018. A 50-Year retrospective on credit risk models, the Altman Z-Score family of models and their applications to financial markets and managerial strategies. Journal of Credit Risk 14, 1–34.

Altman, E., 2013. Predicting financial distress of companies: revisiting the Z-score and Zeta models. Handbook of research methods and applications in empirical finance 428–456.

Altman, E., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance 23, 583–609.

Altman, E., Sabato, G., 2007. Modeling credit risk for SMEs: evidence from the US market. Abacus 43, 332–357.

Altman, E., Sabato, G., Esentato, M., 2020. Assessing the credit worthiness of Italian SMEs and mini-bond issuers. Global Finance Journal 43, 1–14.

Altman, E., Sabato, G., Wilson, N., 2010. The value of non financial information in small and medium-sized enterprise risk management. The Journal of Credit Risk 6, 95–127.

Baixauli, J.S., Modica-Miko, A., 2010. The bias of unhealthy SMEs in bankruptcy prediction models. Journal of Small Business and Enterprise Development 17, 60–77.

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

Barniv, R., Agarwal, A., Leach, R., 2002. Predicting bankruptcy resolution. Journal of Business Finance & Accounting 29, 497–520.

Barro, R.J., McCleary, R., 2003. Religion and economic growth across countries. American Sociological Review 68, 760–781.

Beaver, W., 1966. Financial ratios as predictors of failures. Journal of Accounting Research 4, 71–111.

Belloc, M., Drago, F., Galbiati, R., 2016. Earthquakes, religion, and transition to self-government in Italian cities. The Quarterly Journal of Economics 131, 1875–1926.

Bhargava, M., Dubelaar, C., Scott, T., 1998. Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance. Journal of retailing and consumer services 5, 105–117.

Blanco Olivier, A., Irimia Dieguez, A., Oliver Alfonso, M., Wilson, N., 2015. Improving bankruptcy prediction in micro-entities by using nonlinear effects and non-financial variables. Czech Journal of Economics and Finance 65, 144–166.

Blum, M., 1974. Failing company discriminant analysis. Journal of Accounting Research 12, 1–25.

Bose, I., Pal, R., 2005. Predicting the survival or failure of click-and-mortar corporations: a knowledge discovery approach. European Journal of Operational Research 174, 959–982.
Bruwer, S., Hamman, W.D., 2006. Company failure in South Africa: classification and prediction by means of recursive portioning. South Africa Journal of Business Management 37, 7–18.
Calabrese, R., Marra, G., Osmetti, S.A., 2016. Bankruptcy prediction of small and medium enterprises using a flexible binary generalized extreme value model. Journal of the Operational Research Society 67, 604–615.
Callejón, A.M., Casado, A.M., Fernandez, M.A., Pelaez, J.L., 2013. A system of insolvency prediction for industrial companies using a financial alternative model with neural networks. International Journal of Computational Intelligence Systems 6, 29–37.
Calli, B.A., Coşkun, E., 2021. A longitudinal systematic review of credit risk assessment and credit default predictors. Sage Open (Original Research) 11, 1–19.
Castillo, J.A., Mora-Valencia A., Perote, J., 2018. Moral hazard and default risk of SMEs with collateralized loans. Finance Research Letters 26, 95–99.
Charitou, A., Neophytou, E., Charalambous, C., 2004. Predicting corporate failure: empirical evidence from the UK. European Accounting Review 13, 465–497.
Chen, J., Damanpour, F., Reilly, R., 2010. Understanding antecedents of new product development speed: a Meta-Analysis. Journal of Operations Management 28, 17–33.
Ciampi, F., Giannozzi, A., Marzi, G., Altman, E., 2021. Rethinking SME default prediction: a systematic literature review and future perspectives. Scientometrics 126, 2141–2188.
Ciampi, F., Gordini, N., 2013. Small enterprise default prediction modeling through artificial neural networks: an empirical analysis of Italian small enterprises. Journal of Small Business Management 51, 23–45.
CMS, 2021. Restructuring and insolvency law in Italy [WWW Document].
Constantini, J., 2009. Effects of bankruptcy procedures on firm restructuring: evidence from Italy. Draft working paper.
De Andres, J., Landaio, M., Lorca, P., 2012. Bankruptcy prediction models based on multinorm analysis: an alternative to accounting ratios. Knowledge-Based Systems 30, 67–77.
De Leonardis, D., Rocci, R., 2008. Assessing the default risk by means of a discrete-time survival analysis approach. Applied Stochastic Models in Business and Industry 24, 291–306.
De Massis, A., Chirico, F., Kotlar, J., Naldi, L., 2014. The temporal evolution of proactiveness in family firms: the horizontal S-Curve hypothesis. Family Business Review 27, 35–50.
Deakin, E., 1972. A discriminant analysis of predictors of business failure. Journal of Accounting Research 10, 167–179.
Di Quirico, R., 2010. Italy and the Global Economic Crisis. Bulletin of Italian Politics 2, 3–19.
Du Jardin, P., 2010. Predicting bankruptcy using neural networks and other classification methods: the influence of variable selection techniques on model accuracy. Neurocomputing 73, 2047–2060.
Duell, N., Steinberg, L., Icenogle, G., Chein, J., Chaudhary, N., Di Giunta, L., Dodge, K.A., Fanti, K.A., Lansford, J., Oburu, P., Pastorelli, C., Skinner, A., Sorbring, E., Tapanya, S., Tirado, L.M.U., Alampay, L.P., Al-Hassan, S.M., Takash, H., Bacchini, D., Chang, L., 2018. Age patterns in risk taking across the world. J Youth Adolescence 47, 1052–1072.
Edmister, R.O., 1972. An empirical test of financial ratio analysis for small business failure prediction. Journal of Financial and Quantitative Analysis 7, 1477–1493.
Elnahas, A., Hassan, K., Ismail, G., 2017. Religion and ratio analysis: towards an Islamic corporate liquidity measure. Emerging Markets Review 30, 42–65.
Farooq, U., Qamar, M.A.J., 2019. Predicting multistage financial distress: reflections on sampling, feature and model selection criteria. Journal of Forecasting 38, 632–648.
Fernando, J.M.R., Li, L., Hou, G., 2020. Financial versus non-financial information for default prediction: evidence from Sri Lanka and the USA. Emerging Markets Finance & Trade 56, 673–692.

Fox, J., Monette, G., 1992. Generalized collinearity diagnostics. Journal of the American Statistical Association 87, 178–183.
Gajendran, R., Harrison, D., 2007. The good, the bad, and the unknown about telecommuting: meta analysis of psycological mediators and individual consequences. Journal of Applied Psychology 92, 1524–1541.
Glass, G.V., 1976. Primary, secondary, and Meta-Analysis of research. Educational Researcher 5, 3–8.
Gordini, N., 2014. A genetic algorithm approach for SMEs bankruptcy prediction: empirical evidence from Italy. Expert Systems with Applications 41, 6433–6445.
Gruszczynski, M., 2019. On unbalanced sampling in bankruptcy prediction. International Journal of Financial Studies 7, 28.
Günther, F., Fritsch, S., 2010. Neuralnet: training of neural networks. The R Journal 2, 30.
Gupta, J., Wilson, N., Gregoriou, A., Healey, J., 2014. The effect of internationalisation on modelling credit risk for SMEs: evidence from UK market. Journal of International Financial Markets, Institutions & Money 31, 397–413.
Hammond, P., Opoku, M.O., Kwakwa, P.A., Amissah, E., 2023. Predictive models of going concerns and business failure. Cogent Business & Management 10, 1–24.
Hand, D.J., 2004. Marginal classifier improvement and reality. Paper presented at the Symposium on Data Mining.
Henningsen, D.M., 2012. The origins of the Italian sovereign debt crisis. CMC Senior Theses Paper 379.
Hu, H., Sathye, M., 2015. Predicting financial distress in the Hong Kong growth enterprises market from the perspective of financial sustainability. Sustainability 7, 1186–1200.
Huang, B., Yao, X., Luo, Y., Li, J., 2023. Improving financial distress prediction using textual sentiment of annual reports. Annals of Operations Research 1, 457–484.
Hunter, J., Schmidt, F., 2004. Methods of meta-analysis: correcting error and bias in research findings, Sage. ed.
Hunter, J., Schmidt, F., 1990. Dichotomization of continuous variables: the implications for meta-analysis. Journal of applied psychology 75, 334.
Kanapickiene, R., Kanapickas, T., Neciunas, A., 2023. Bankruptcy prediction for micro and small enterprises using financial, non-financial, business sector and macroeconomic variables: the case of the Lithuanian construction sector. Risks 11, 1–33.
Kim, H., Cho, H., Ryu, D., 2020. Corporate default predictions using machine learning: literature Review. Sustainability 12, 1–11.
Kuizinienė, D., Krilavičius, T., Damaševičius, R., Maskeliūnas, R., 2022. Systematic review of financial distress identification using artificial intelligence methods. Applied Artificial Intelligence 36, 3571.
Laitinen, E.K., 2010. Financial variables in payment default prediction: test of iteration effects. International review of applied financial Issues and Economics 2, 98–123.
Laitinen, E.K., Camacho-Miñano, M., Muñoz-Izquierdo, N., 2023. A review of the limitations of financial failure prediction research. Spanish accounting review 26, 255–273.
Laitinen, E.K., Laitinen, T., 2000. Bankruptcy prediction: application of the Taylor's expansion in logistic regression. International Review of Financial Analysis 9, 327–349.
Lepore, A., 2020. L'evoluzione del divario tra il Nord e il Sud dal dopoguerra a oggi. Estudios Historicos 12, 1–63.
Li, X., Wang, K., Liu, L., Xin, J., Yang, H., Gao, C., 2011. Application of the entropy weight and TOPSIS method in safety evaluation of coal mines. Procedia Engineering 26, 2085–2091.
Lin, L., Piesse, J., 2004. Identification of corporate distress in UK industrials: a conditional probability analysis approach. Applied Financial Economics 14, 73–82.
Liou, D., Smith, M., 2007. Macroeconomic variables and financial distress. Journal of Accounting – Business & Management 14, 17–31.

Mata, R., Josef, A.K., Hertwig, R., 2016. Propensity for risk taking across the life span and around the globe. Psychological Science 27, 231–243.
McLeay, S., Omar, A., 2000. The sensitivity of prediction models to the non-normality of bounded and unbounded ratios. British Accounting Review 32, 213–230.
Mokrišová, M., Horváthová, J., 2023. Domain knowledge features versus LASSO features in predicting risk of corporate bankruptcy—DEA approach. Risks 11, 1–18.
Mućko, P., Adamczyk, A., 2023. Does the bankrupt cheat? Impact of accounting manipulations on the effectiveness of a bankruptcy prediction. PLOS ONE 18, 1–13.
Mujkić, E., Poljašević, J., 2023. Prediction of insolvency using logistic regression: the case of the Republic of Srpska. Ekonomski vjesnik/Econviews - Review of Contemporary Business, Entrepreneurship and Economic Issues 36, 127–141.
Myoung-Jong, K., Dae-Ki, K., 2010. Ensemble with neural networks for bankruptcy prediction. Expert Systems with Applications 37, 3373–3379.
Naldi, L., Nordqvist, M., Sjöberg, K., Wiklund, J., 2007. Entrepreneurial orientation, risk taking, and performance in family firms. Family Business Review 20, 33–47.
Omelka, J., Beranová, M., Tabas, J., 2013. Comparison of the models of financial distress prediction. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis 61, 2587–2592.
Park, S., Kim, H., Kwon, J., Kim, T., 2021. Empirics of Korean shipping companies' default predictions. Risks 9, 1–17.
Pasetto, L., n.d. Stato di crisi e insolvenza: i fondamenti della disciplina concorsuale italiana.
Pindado, J., Rodriguez, L., 2004. Parsimonious model of financial insolvency in small companies. Small Business Economics 22, 51–66.
Platt, H., Platt, M., 2002. Predicting corporate financial distress: reflection on choice based sample bias. Journal of Economics and Finance 26, 184–199.
Pratt, M.J., Van Peursem, K., 2002. A New Zealand failure prediction model: development and international implications. Advances in International Accounting 15, 229–247.
Realization of entropy weights method in R [WWW Document], n.d. . Programmer Sought.
Resti, A., Sironi, A., 2007. Risk management and shareholders' value in banking, Wiley Finance Series.
Rose, P., Giroux, G.A., 1984. Predicting corporate bankruptcy: an analytical and empirical evaluation. Review of Business and Economic Research 19, 1–12.
Sabato, G., 2008. Managing credit risk for retail low-default portfolios, in: Credit Risk. pp. 293–312.
Sayari, N., Simga Mygan, K., 2017. Industry specific financial distress modeling. Business Research Quarterly 20, 45–62.
Scherger, V., Terceno, A., Vigier, H.P., 2019. A systematic overview of the prediction of business failure. International Journal of Technology Policy and Management 19, 196–211.
Shi, Y., Li, X., 2019. An overview of bankruptcy prediction models for corporate firms: a systematic literature review. Intangible Capital 15, 114–127.
Shumway, T., 2001. Forecasting bankruptcy more accurately: a simple hazard model. The Journal of Business 74, 101–124.
Sirirattanaphonkun, W., Pattarathammas, S., 2012. Default prediction for small-medium enterprises in emerging market: evidence from Thailand. Seoul Journal of Business 18, 25–54.
Skogsvik, K., Skogsvik, S., 2013. On the choice based sample bias in probabilistic bankruptcy prediction. Investment Management and Financial Innovations 10, 29–37.
Souma, W., Vodenska, J., Aoyama, H., 2019. Enhanced news sentiment analysis using deep learning methods. Journal of Computational Social Science 2, 33–46.
Stanley, T.D., Jarrell, S., 1989. Meta-regression analysis: a quantitative method of literature surveys. Journal of Economic surveys 3, 299–308.

Stearns, S., 2000. Life history evolution: successes, limitations, and prospects. Naturwissenschaften 87, 476–486.
Stevenson, M., Mues, C., Bravo, C., 2021. The value of text for small business default prediction: a deep learning approach. European Journal of Operational Research 295, 758–771.
Tai, B., Tai, L., 1986. A multivariate statistical analysis of the characteristics of problem Firms in Hong Kong. Asia Pacific Journal of Management 3, 121–127.
Tinoco, M.H., Wilson, N., 2013. Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. International Review of Financial Analysis 30, 394–419.
Tseng, F., Hu, Y., 2009. Comparing four bankruptcy prediction models: logit, quadratic interval logit, neural and fuzzy neural networks. Expert Systems with Applications 37, 1846–1853.
Valaskova, K., Gajdosikova, D., Belas, J., 2023. Bankruptcy prediction in the post-pandemic period: a case study of Visegrad Group countries. Oeconomia Copernicana 14, 253–293.
Veganzones, D., Severin, E., 2021. Corporate failure prediction models in the twenty-first century: a review. European Business Review 33, 204–226.
Vu, N.T., Nguyen, N.H., Tran, T., Le, B.T., Vo, D.H., 2023. A LASSO-based model for financial distress of the Vietnamese listed firms: does the covid-19 pandemic matter? Cogent Economics & Finance 11, 1–17.
Wald, K.D., Owen, D.E., Hill, S., 1990. Political cohesion in churches. The Journal of Politics 52, 197–215.
Wang, K., Wan, J., Li, G., Sun, H., 2022. A hybrid algorithm-level ensemble model for imbalanced credit default prediction in the energy industry. Energies 15, 1–18.
Wilson, N., Chong, K., Peel, M., 1995. Neural network simulation and the prediction of corporate outcomes: some empirical findings. International Journal of the Economics of Business 2, 31–50.
Yazdanfar, D., 2011. Predicting bankruptcy among SMEs: evidence from Swedish firm-level data. International Journal of Entrepreneurship and Small Business 14, 551–565.
Yoo, W., Mayberry, R., Bae, S., Singh, K., Qinghua He, P., Lillard, J.W., 2014. A study of effects of multicollinearity in the multivariable analysis. International Journal of Applied Science and Technology 4, 9–19.
Zhao, Y., Lin, D., 2023. Prediction of micro- and small-sized enterprise default risk based on a logistic model: evidence from a bank of China. Sustainability 15, 1–13.
Zizi, Y., Jamali-Alaoui, A., El Goumi, B., Oudgou, M., El Moudden, A., 2021. An optimal model of financial distress prediction: a comparative study between neural networks and logistic regression. Risks 9, 1–24.
Zmijewski, M., 1984. Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research 22, 59–82.
Zuckerman, M., 2007. The sensation seeking scale V (SSS-V): still reliable and valid. Personality and Individual Differences 43, 1303–1305.

ANNEX – TABLE 5: LITERATURE SELECTION SUMMARY – UNDERLYING DATA

Reference	Year	Sample Size	% Bankruptcy	Random sampling technique	Type of Firm	Accounting variables only	Modelling technique	Type I error rate (in sample)	Type II error rate (in sample)	Type I error rate (out of sample)	Type II error rate (out of sample)
E. Altman	1968	<100	50%	No	Listed	No	MDA	94%	97%	96%	79%
E. Deakin	1972	<100	50%	No	Listed	Yes	MDA	86%	87%	n.a.	n.a.
R. Edmister	1972	<100	50%	No	Mixed / Other	Yes	MDA	90%	95%	n.a.	n.a.
Rose S.and Giroux P.	1984	<100	50%	No	Mixed / Other	No	MDA	85%	97%	n.a.	n.a.
B. Y.K Tai and L. S.T. Tai	1986	<100	50%	No	Listed	No	MDA	100%	82%	91%	73%
Sung T., Chang N. and Lee G.	1999	<10,000	36%	Yes	Listed	No	MDA	69%	90%	n.a.	n.a.
Sung T., Chang N. and Lee G.	1999	<10,000	36%	Yes	Mixed / Other	No	Other	72%	90%	n.a.	n.a.
E.K. Laitinen	2000	<10,000	50%	No	Mixed / Other	Yes	CPM	82%	80%	74%	75%
Sirirattanaphonkun W.	2000	<10,000	33%	Yes	Mixed / Other	Yes	MDA	52%	84%	62%	90%
Sirirattanaphonkun W.	2000	<10,000	33%	Yes	Mixed / Other	Yes	CPM	55%	99%	58%	99%
Sirirattanaphonkun W.	2000	<10,000	33%	Yes	Mixed / Other	Yes	CPM	58%	95%	58%	96%
N. Wilson	1995	<10,000	36%	Yes	Listed	No	CPM	93%	97%	n.a.	n.a.
N. Wilson	1995	<10,000	36%	Yes	Listed	No	NN	98%	100%	95%	95%
Anandarajan M.	2001	<10,000	20%	Yes	Listed	No	NN	96%	87%	94%	70%
Anandarajan M.	2001	<10,000	20%	Yes	Listed	No	NN	99%	99%	95%	94%
Charalambous C., Charitou A. and Kaorou F.	2001	<10,000	50%	No	Listed	No	CPM	70%	86%	n.a.	n.a.
Charalambous C., Charitou A. and Kaorou F.	2001	<10,000	50%	No	Listed	No	NN	88%	77%	n.a.	n.a.
K. A. Van Peursem and M.J. Pratt	2002	<10,000	50%	No	Listed	Yes	CPM	92%	92%	88%	75%
A. Charitou et al	2004	<10,000	50%	No	Listed	No	CPM	92%	96%	86%	76%
A. Charitou et al	2004	<10,000	50%	No	Listed	No	NN	100%	92%	90%	76%
A. Charitou et al	2004	<10,000	50%	No	Listed	No	CPM	84%	87%	82%	82%
A. Charitou et al	2004	<10,000	50%	No	Listed	No	MDA	92%	96%	90%	75%
L. Lin and J. Piesse	2004	<100	42%	No	Mixed / Other	No	CPM	78%	91%	67%	75%
J. Pindado, L.F. Rodriguez	2004	<10,000	50%	No	Mixed / Other	Yes	CPM	92%	92%	88%	76%
Steyn Bruwer	2006	<10,000	40%	Yes	Mixed / Other	No	Other	67%	80%	61%	78%
I. Bose, R. Pal	2006	<10,000	50%	No	Mixed / Other	Yes	MDA	70%	73%	69%	67%
I. Bose, R. Pal	2006	<10,000	50%	No	Mixed / Other	Yes	NN	73%	84%	67%	79%
I. Bose, R. Pal	2006	<10,000	50%	No	Mixed / Other	No	NN	82%	88%	72%	77%
G. Sabato	2006	>10,000	6%	Yes	Mixed / Other	No	CPM	81%	73%	n.a.	n.a.
G. Sabato	2006	>10,000	6%	Yes	Mixed / Other	No	CPM	85%	76%	n.a.	n.a.
E. Altman, G. Sabato	2007	<10,000	6%	Yes	Mixed / Other	Yes	CPM	n.a.	n.a.	91%	75%
E. Altman, G. Sabato	2007	<10,000	6%	Yes	Mixed / Other	Yes	CPM	n.a.	n.a.	80%	72%
E. Altman, G. Sabato	2007	<10,000	6%	Yes	Mixed / Other	Yes	MDA	n.a.	n.a.	74%	70%
Dah-Kwei Liou and Malcolm Smith	2007	<10,000	18%	Yes	Listed	No	#N/A	n.a.	n.a.	81%	n.a.
F. Tseng	2009	<100	71%	No	Mixed / Other	Yes	CPM	77%	n.a.	86%	n.a.

Reference	Year	Sample Size	% Bankruptcy	Random sampling technique	Type of Firm	Accounting variables only	Modelling technique	Type I error rate (in sample)	Type II error rate (in sample)	Type I error rate (out of sample)	Type II error rate (out of sample)
F. Tseng	2009	<100	71%	No	Mixed / Other	Yes	NN	91%	n.a.	95%	n.a.
E. Altman, G. Sabato	2010	>10,000	1%	Yes	Mixed / Other	Yes	CPM	76%	75%	76%	73%
E. Altman, G. Sabato	2010	>10,000	1%	Yes	Mixed / Other	Yes	CPM	80%	76%	77%	73%
E.K. Laitinen	2010	>10,000	3%	Yes	Mixed / Other	Yes	CPM	74%	74%	73%	74%
M.H. Tinoco	2013	<10,000	13%	Yes	Listed	No	CPM	87%	85%	n.a.	n.a.
M.H. Tinoco	2013	<10,000	13%	Yes	Listed	No	CPM	82%	80%	n.a.	n.a.
M.H. Tinoco	2013	<10,000	13%	Yes	Listed	No	NN	88%	85%	n.a.	n.a.
Hu H. and Sathye M.	2015	<10,000	30%	Yes	Listed	No	CPM	n.a.	n.a.	75%	91%
Jabeur S. and Fahmi Y.	2015	<10,000	50%	No	Listed	No	MDA	97%	94%	n.a.	n.a.
Jabeur S. and Fahmi Y.	2015	<10,000	50%	No	Listed	No	CPM	77%	70%	n.a.	n.a.
Jabeur S. and Fahmi Y.	2015	<10,000	50%	No	Listed	No	Other	97%	97%	n.a.	n.a.
Blanco Olivier A., Irimia Dieguez A., Oliver Alfonso M. and Wilson N.	2015	>10,000	2%	Yes	Mixed / Other	No	CPM	75%	71%	75%	71%
Blanco Olivier A., Irimia Dieguez A., Oliver Alfonso M. and Wilson N.	2015	>10,000	2%	Yes	Mixed / Other	No	NN	77%	74%	76%	74%
Nouri B. and Soltani M.	2015	<100	23%	Yes	Listed	No	CPM	74%	96%	n.a.	n.a.
A.M. Elnahas M.K. Hassan	2016	<10,000	16%	Yes	Listed	No	MDA	n.a.	n.a.	84%	69%
A.M. Elnahas M.K. Hassan	2016	<10,000	16%	Yes	Listed	No	MDA	n.a.	n.a.	88%	72%
Mingjing Wang a, HuilingChen a,n, HuaizhongLi b, ZhennaoCai a, XuehuaZhao c, Changfei Tong a, JunLi a, XinXu	2016	<10,000	47%	No	Mixed / Other	Yes	NN	90%	82%	n.a.	n.a.
Mingjing Wang a, HuilingChen a,n, HuaizhongLi b, ZhennaoCai a, XuehuaZhao c, Changfei Tong a, JunLi a, XinXu	2016	<10,000	50%	No	Mixed / Other	Yes	NN	88%	86%	n.a.	n.a.
Waqas Hamid and Md-Rus Rohani	2018	<10,000	16%	Yes	Listed	No	CPM	56%	99%	59%	99%
Ma'aji M., Hiau A. and Khaw K.	2019	<10,000	50%	No	Mixed / Other	No	MDA	95%	86%	100%	82%
Ouenniche J., Bouslah K., Gladish B. and Xu B.	2019	<10,000	6%	Yes	Listed	Yes	Other	100%	100%	100%	100%
Khademolqorani, Ali Zeinal Hamadani, and FarimahMokhatab Rafiei	2015	<10,000	32%	Yes	Listed	Yes	MDA	86%	57%	n.a.	n.a.
Khademolqorani, Ali Zeinal Hamadani, and FarimahMokhatab Rafiei	2015	<10,000	32%	Yes	Listed	Yes	CPM	89%	57%	n.a.	n.a.
Khademolqorani, Ali Zeinal Hamadani, and FarimahMokhatab Rafiei	2015	<10,000	32%	Yes	Listed	Yes	NN	95%	91%	n.a.	n.a.
Khademolqorani, Ali Zeinal Hamadani, and FarimahMokhatab Rafiei	2015	<10,000	32%	Yes	Listed	Yes	Other	94%	95%	n.a.	n.a.
Gordini N.	2014	<10,000	42%	No	Mixed / Other	Yes	Other	80%	70%	n.a.	n.a.
Gordini N.	2014	<10,000	42%	No	Mixed / Other	Yes	NN	79%	68%	n.a.	n.a.
Gordini N.	2014	<10,000	42%	No	Mixed / Other	Yes	CPM	78%	60%	n.a.	n.a.
Qi Yu, Yoan Miche, Eric Séverin, Amaury Lendasse	2013	<10,000	50%	No	Listed	Yes	NN	92%	97%	n.a.	n.a.
Tsai C.	2011	<10,000	56%	No	Mixed / Other	No	NN	89%	93%	100%	47%
Tsai C.	2011	<10,000	30%	Yes	Mixed / Other	No	NN	95%	76%	80%	95%
Tsai C.	2011	<10,000	55%	No	Mixed / Other	No	NN	94%	95%	93%	98%

Reference	Year	Sample Size	% Bankruptcy	Random sampling technique	Type of Firm	Accounting variables only	Modelling technique	Type I error rate (in sample)	Type II error rate (in sample)	Type I error rate (out of sample)	Type II error rate (out of sample)
Tsai C.	2011	<10,000	46%	No	Mixed / Other	No	NN	90%	91%	100%	95%
Tsai C.	2011	>10,000	50%	No	Mixed / Other	No	NN	86%	88%	74%	71%
Ahmad Ahmadpour Kasgari, Mehdi Divsalar, Mohamad Reza Javid, Seyyed Javad Ebrahimian	2012	<10,000	48%	No	Listed	Yes	NN	93%	96%	95%	93%
Ahmad Ahmadpour Kasgari, Mehdi Divsalar, Mohamad Reza Javid, Seyyed Javad Ebrahimian	2012	<10,000	48%	No	Listed	Yes	CPM	74%	80%	91%	75%
Callejón A. M., Casado A. M, Fernández M. A., Peláez J.I.	2013	<10,000	50%	No	Mixed / Other	Yes	NN	95%	90%	95%	90%
Zhoua L., Keung K., Yend J.	2012	<100	50%	No	Mixed / Other	Yes	MDA	67%	64%	n.a.	n.a.
Zhoua L., Keung K., Yend J.	2012	<100	50%	No	Mixed / Other	Yes	CPM	87%	60%	n.a.	n.a.
Zhoua L., Keung K., Yend J.	2012	<100	50%	No	Mixed / Other	Yes	NN	80%	69%	n.a.	n.a.
Zhoua L., Keung K., Yend J.	2012	<10,000	50%	No	Mixed / Other	Yes	MDA	99%	64%	n.a.	n.a.
Zhoua L., Keung K., Yend J.	2012	<10,000	50%	No	Mixed / Other	Yes	CPM	97%	96%	n.a.	n.a.
Zhoua L., Keung K., Yend J.	2012	<10,000	50%	No	Mixed / Other	Yes	NN	96%	96%	n.a.	n.a.
Ning Chen, Bernardete Ribeiro, Armando S. Vieira, João Duarte, João C. Neves	2011	<10,000	50%	No	Mixed / Other	Yes	NN	89%	95%	83%	89%
Mehdi Divsalar, Ali Khatami Firouzabadi, Meisam Sadeghi, Amir Hossein Behrooz and Amir Hossein Alavi	2011	<10,000	49%	No	Listed	No	NN	100%	98%	96%	94%
De Andrés J., Landajo M., Lorca P.	2011	<10,000	50%	No	Mixed / Other	No	MDA	76%	80%	75%	71%
De Andrés J., Landajo M., Lorca P.	2011	<10,000	50%	No	Mixed / Other	No	CPM	79%	82%	74%	79%
De Andrés J., Landajo M., Lorca P.	2011	<10,000	50%	No	Mixed / Other	No	NN	82%	86%	73%	78%
Du Jardin P.	2010	<10,000	50%	No	Mixed / Other	Yes	CPM	90%	91%	90%	94%
Du Jardin P.	2010	<10,000	50%	No	Mixed / Other	Yes	NN	92%	93%	95%	92%
Kim M., Kang D.	2009	<10,000	50%	No	Listed	Yes	NN	82%	68%	n.a.	n.a.
Fernando, Li, Hou	2020	<10,000	50%	No	Listed	No	CPM	84%	n.a.	n.a.	n.a.
Fernando, Li, Hou	2020	<10,000	50%	No	Listed	No	CPM	91%	n.a.	n.a.	n.a.
Stevenson, Mues, Bravo	2021	>10,000	20%	Yes	Mixed / Other	No	CPM	82%	n.a.	76%	n.a.
Stevenson, Mues, Bravo	2021	>10,000	20%	Yes	Mixed / Other	No	Other	83%	n.a.	74%	n.a.
Stevenson, Mues, Bravo	2021	>10,000	20%	Yes	Mixed / Other	No	Other	89%	n.a.	76%	n.a.
Park, Kim, Kwon and Kim	2021	<10,000	13%	Yes	Mixed / Other	Yes	CPM	86%	61%	n.a.	n.a.
Wang, Wan, Li, Sun	2022	<10,000	6%	Yes	Mixed / Other	Yes	CPM	65%	96%	n.a.	n.a.
Wang, Wan, Li, Sun	2022	<10,000	6%	Yes	Mixed / Other	Yes	NN	54%	98%	n.a.	n.a.
Wang, Wan, Li, Sun	2022	<10,000	6%	Yes	Mixed / Other	Yes	Other	91%	89%	n.a.	n.a.
Zizi, Jamali-Alaoui, El Goumi, Oudgou, El Moudden	2021	<10,000	46%	Yes	Mixed / Other	Yes	CPM	93%	97%	n.a.	n.a.
Zizi, Jamali-Alaoui, El Goumi, Oudgou, El Moudden	2021	<10,000	46%	Yes	Mixed / Other	Yes	NN	87%	90%	n.a.	n.a.

LIST OF FIGURES

LIST OF TABLES