Rating, capital structure and bankruptcy prediction

Michele Modina, Filomena Pietrovito

Università degli Studi del Molise, Campobasso, Italy

Abstract

Small and medium-sized firms (SMEs) play a fundamental role into the economy of many countries, including Italy. The simpleness of the financial structure of Italian SMEs, relying mostly on bank loans, is a constraint towards growth whereas the limited variety of financing instruments restricts the investment capacity and their competitiveness. Building on this framework, this paper aims to analyze the determinants of the default probability for a sample of 9,208 Italian limited liabilities SMEs in a time frame of three years, over the period 2006-2010. Specifically, this paper adopts a logistic regression model constructed from data provided by the *Centrale Rischi Finanziari* (CRIF), an Italian credit rating agency. The results show that, among the observable financial and economic characteristics of firms, the capital structure (both in terms of internal and external funds and in terms of the source of external financing) and interest expenses are more relevant than economic variables as determinants of SMEs' default.

Jel Classification: G21; D21

Keywords: default prediction, capital structure, small and medium-sized enterprises, logistic regression

M. Modina (corresponding author), Università degli Studi del Molise, Dipartimento di Economia, Gestione, Società e Istituzioni, II Edificio Polifunzionale, Via F. de Sanctis, 86100 Campobasso, phone: (++) 39 3338447284 e-mail: michele.modina@unimol.it

F. Pietrovito, Università degli Studi del Molise, Dipartimento di Economia, Gestione, Società e Istituzioni, II Edificio Polifunzionale, Via F. de Sanctis, 86100 Campobasso, phone: (++) 39 0874 404335 e-mail: filomena.pietrovito@unimol.it

1. Introduction

Small and medium sized firms (SMEs with fewer than 250 employees) constitute the backbone of the economy of many countries (Altman and Sabato 2007). They account for over 95% and up to 99% of the business populations whose contribution varies considerably across countries (OECD 2005). In the European economy, for example, they account for 99.8 per cent of non-financial firms (91.2% are micro-enterprises, with fewer than ten employees), corresponding to 20.7 million businesses. In terms of the number of employees, SMEs provide 67.4% of jobs in the non-financial business economy. The relevance of SMEs is particularly significant for some UE members, including Italy, where small firms (with fewer than 50 employees) constitute 99% of manufacturing enterprises (OECD 2005).

As argued by Pederzoli and Torricelli (2010), one of the main issue for Italian SMEs is to recover money to finance their investment. Banks play an important role in this framework for two prime reasons: (*i*) they are the only subject issuing loans directly to SMEs and (*ii*) other financial sources, such as issuing shares on capital markets, are barely used by Italian SMEs. However, when the stock of debt is about to saturate, the firm's growth is plumped by self-financings. Consequently, the simpleness of the financial structure of Italian SMEs is a constraint towards growth whereas the limited variety of financing instruments restricts the investment capacity and, therefore, their competitiveness. Looking at it from the credit risk, Italian SMEs are therefore riskier than larger counterparts.

However, the Basel II and the Basel III regulations allow banks to adopt an internal rating system (IRS) approach and require to pay attention to measuring and managing credit risk in order to evaluate the obligor's probability of default. Basel II imposes banks minimum capital requirements which are based on the probability of default and which recognize a different treatment for exposure towards SMEs (Basel Committee on Banking Supervision 1999, 2004). As a consequence, banks need to develop credit risk models specifically addressed to SMEs in order to better manage their lending activity (Rikkers and Thibeault 2009).

In the path for such framework, this paper focuses on the role that financial and economic factors, drawn from balance sheets and income statement, can play in affecting the probability of default. The work analyzes the determinants of default for a large sample of 9,208 limited liabilities Italian SMEs in a time frame of three years, over the period 2006-2010 that partially includes the effects of the recent financial crisis. Data are provided by the *Centrale Rischi Finanziari* (CRIF), an Italian rating agency. The definition of SMEs reflects the same provided by the Basel II regulation, and confirmed in Basel III. Similarly, the default definition includes suffering, 180 days past-due loans (instead of 90-days, following the past-due trigger provided by the Italian supervisor), bankruptcy and ascertained losses.

Since the focus is to detect variables predicting the default event after three years, the sample includes both default and non-default firms in 2009-2010. The analysis is conducted on the 2006 (2007) balance sheet and income statement data for firms in default in 2009 (2010) and balance sheet and income statement data in 2006 for non-default firms. The time period examined also includes the effect of the recent global crisis since we analyse the determinants of the probability of default in 2009 and 2010, that is, one and two years *after* the crisis. In practice, we take into account the characteristics of the firm before the crisis in order to evaluate their effect after that.[1]

By applying factor analysis – to reduce the original set of information – and the stepwise logistic regression – for the default prediction – the results show important links between the financial structure and the probability of default. Specifically, these indicators include the level of capitalization, interest expenses and bank loans.

Compared to the existing literature, this paper offers a quadruple contribution. *First*, while typical studies on firm default focus on large publicly traded firms (Bottazzi et al. 2011), our study examines a sample of limited liabilities firms representing the universe of Italian SMEs operating in the manufacturing, commercial and service sectors. We use balance sheets and income statement data that are publicly available, contrarily to market data, and that can be easily accessed by banks. *Second*, unlike the existing studies on Italian firms, this paper is more 'forward-looking' since it extends the prediction time to three years (instead of one), mitigating the influence of the economic cycle on the probability of default.[2] *Third*, this paper adopts a broad default definition that not only includes "bankruptcy" or "liquidation". These two legal definitions are indeed too much restrictive the very moment that a legal failure does not reflect the real failure event (Rikkers and Thibeault 2009). *Fourth*, we focus on indicators concerning the financial structure of firms, by including specific capital structure indicators and several indicators of the interest expenses, along to traditional ones. A reliable and accurate default model suitable to SMEs should indeed consider the key ratio around capital structure in order to

---

[1] However, since we adopt a three-year window, with available data, we cannot distinguish between the effects of these characteristics before and after the financial crisis.

[2] An exception is provided by Ciampi and Gordini (2013a) adopting a lag time of four years and by Ciampi and Gordini (2013b) that estimate the default probability in a 3-years window, but only for the period 2005-2008.

assess successfully the credit risk of SMEs and implement an appropriate default prediction process.

From a policy implications mere point of view, this paper highlights the close relationship between the capital structure of Italian SMEs and their probability of default. A high debt (especially the bank one) and the high incidence of interest expenses are significant factors in anticipating insolvency. Therefore, our findings suggest that: (*i*) financial institutions should include financial structure variables, and, above all, those variables concerning both the composition of sources of funding and the interest expenses related to default prediction models suitable to SMEs; and (*ii*) Italian firms need to adopt new approaches and new financing instruments to strengthen their funding policies.

The paper has been subdived into six sections. Section 2 presents a review of the literature concerning the most important statistical techniques applied for the default prediction and the Italian SMEs. Section 3 describes the data set and the main statistics. Section 4 shows the methodology and the results of empirical analysis. Section 5 illustrates the robustness checks, including the model validation. Section 6 draws some conclusions.

2. Literature review

The vast literature on the default prediction models assigns a probability of failure to firms, in a given time horizon, by looking at financial and economic indicators. This strand of the literature proposes several techniques to deal with the issue (Rikkers and Thibeault 2009; Pederzoli and Torricelli 2010). The first group of techniques includes *structural form models* (Merton 1974) and *reduced form models* (Jarrow and Turnbull 1995) exploitable with market data on stocks or corporate bonds and asset swaps. These models do not fit (not listed) SMEs since they require market data. The second group includes *traditional models* (Beaver 1966; Altman 1968; Ohlson 1980) using accounting data which are available for all firms, regardless of their size. The accounting-based traditional models are indeed the most suitable for privately-held firms with no market data available. Traditional methods include both intelligent techniques (such as, expert systems and neural network) and statistical methods (multivariate discriminant analysis, logit and probit model).[3]

[3] For a comprehensive review of the literature, see e.g. Ravi Kumar and Ravi (2007).

This section reviews the most influential contributions, focusing on the statistical techniques used to predict default probability. Moreover, it analyzes the works investigating default prediction models specifically developed for SMEs.

2.1 Statistical techniques used in the literature

The main statistical techniques used to solve the default prediction problems for firms can be sorted into four categories (Ravi Kumar and Ravi 2007): (*i*) the linear discriminant analysis (LDA); (*ii*) the multivariate discriminant analysis (MDA); (*iii*) the logistic regression (LR); and (*iv*) the probit regression (PR).

The first two models are based on the *a priori* assumption that there are two mutually exclusive groups of firms (default and non-default) and that the differences between them can be captured by observing (individually) a set of financial ratios (Foglia et al. 1998). On the other hand, the more recent logistic and probit approaches aim to assign each firm its probability of default on the basis of several financial indicators.

The seminal paper by Beaver (1966) introduced the LDA that individually examines various financial indicators to capture those factors capable to explain the current status of the firm and its future perspectives. By using a matched sample (including 79 failed and 79 non-failed firms), Beaver (1966) analyzes the predictive power of a set of financial indicators. More specifically, he uses a dichotomous classification test to identify an appropriate cut-off point for each indicator that is likely to minimize the number of errors made in classifying firms as failed and non-failed. The study finds that the cash flow to total debts ratio shows the best performance.

However, as pointed out in the literature, one of the major shortcomings of the LDA is that it considers indicators, individually or in a system, without attempting to combine them in a synthetic quantitative measure (Varetto 1999). In a more influential contribution, Altman (1968) successfully applies the MDA in order to solve the inconsistency problem related to Beaver's approach and to explore a more complete financial profile of firms (Altman and Sabato 2007). Altman (1968) examines a matched sample including 33 failed and 33 non-failed firms that experienced bankruptcy over the period 1946-1965. He finds that liquidity, profitability, leverage, solvency and activity ratios do a good job in discriminating between the two groups of failed and non-failed firms. Starting from the pioneering study of Altman (1968), several studies have been carried out on the firm default prediction modelling focusing on different samples and groups of indicators (Eisenbeis 1977; Piesse and Wood 1992; Altman et al. 1994; Foglia et al. 1998; Grice and Ingram 2001).

While recognizing a significant maturity to the MDA, this statistical technique requires restrictive assumptions, especially on independent variables (Altman and Sabato 2007): the financial indicators used are multivariate normally distributed, thus preventing the use of dummy variables; and the variance-covariance matrix of financial indicators are equal in the groups of failed and non-failed firms.

To deal with these issues, more recent studies, starting from the seminal paper by Ohlson (1980), adopt the LR analysis to predict the default probability of firms. The advantages of this approach rely on the following issues: missing assumptions of MDA on independent variables; the coefficients of the independent variables are interpreted as their own effects on the probability of default; it allows to handle with both quantitative and qualitative independent variables; it affords to work with non-matched samples; averting a discrimination between default and non-default firms, as in the discriminant analysis, logistic regression defines a ranking in firms classification; and currently, it is the most common used methodology by bank credit risk system (Altman and Sabato 2007; Dainelli et al. 2013).

Regarding the empirical applications of the LR, Ohlson (1980) studies the impact of nine different financial indicators on the probability of default of 105 default and more than 2,000 non-default firms during the period 1970-1976. Through this methodology, he identifies four categories of factors that influence the likelihood of firm default: firm size, financial structure indicators, performance indicators and current liquidity. Following this seminal paper, recent literature enriched by numerous studies using the LR (Platt and Platt 1990; Laviola and Trapanese 1997; Mossman et al. 1998; Becchetti and Sierra 2003; Altman and Sabato 2007; Pierri et al. 2011). In spite of, from a theoretical point of view, MDA and LR are different, empirical analyses show that they provide quite similar results in terms of prediction accuracy (Altman and Sabato 2007).

On a related ground, Bottazzi et al. (2011) estimate a PR to predict the default probability for a large sample of limited liabilities Italian firms. This paper verifies the significance of financial ratios, also controlling economic ratios, and finds that among financial variables the cost of debts exerts the most important impact on the default probability. Concerning the economic variables, profitability and productivity reduce the likelihood of default, whereas firm size exerts a positive impact. Moreover, economic variables in their framework play a short and long term effect on the probability of default.

However, logit and probit techniques work with monotonic relationships between dependent variables (default/non-default) and independent variables (economic–financial ratios). Therefore, new innovative methodologies taken from other disciplines, such as neural networks and genetic algorithms, have been taken in order to relax the requirements on data and/or lowers the dependence on heuristics (Härdle et al. 2009).[4]

2.2 SMEs studies

As highlighted in Section 1, Basel II imposes banks minimum capital requirements based on the firms probability of default and which recognize a different treatment for exposure towards SMEs (Pederzoli and Torricelli 2010). For this reason, many analysts have focused on the SMEs segment. However, many of them concentrate on the impact of Basel II on bank capital requirement and thus on possible effects on credit rationing (Altman and Sabato 2005, 2007). A related strand of the literature states indeed that, even though SMEs are the engine of economic development, they are more financially constrained then large firms and access to finance is an important growth constraint for them. Indeed, market and institutional failures impede their growth, justifying government interventions (Giannetti 2003; Beck and Demirgüc-Kunt 2006; Demirgüc-Kunt, Detragiache and Tressel 2008). Moreover, innovative financing instruments can help facilitate SMEs' access to finance even in the absence of well-developed institutions.

Until Basel II, credit risk modeling expressly designed for SMEs received marginal attention, albeit in the last years the literature on this issue is increasing. As a result, a significant number of studies aims at analyzing and predicting the bankruptcy risk of SMEs in different geographical contexts.

For example, Altman and Sabato (2005) develop three different models to assign the probability of default to firms in Australia, US and Italy. They specifically consider the differences between the structures of SMEs and their credit risk attributes in different countries. In a more recent study focused on US, Altman and Sabato (2007) use a logit regression technique to develop a one year default prediction model on a panel of 2,000 firms over the period 1994-2002. They demonstrate that a convenient way for banks in setting their credit risk system is to separate SMEs from generic firms. They find that such models properly directed to SMEs should contain leverage, liquidity, profitability, coverage and activity indicators.

---

[4] As far as the use of different statistical methods is concerned, the relationship between generalisation and specificity in these models is another widely debated theme. Several authors are discussing about the need to develop industry-specific models rather general ones. In most cases, industry-specific model oriented researches (Altman 1994; Sironi 2003) are more accurate than general models. This is probably due to greater homogeneity of financial indicators within specific industries.

Nevertheless, for a sample of 310 German firms, Rikkers and Thibeault (2009) develop a logit model including standard financial ratios and a structural form measure. They find a variable with an additional predictive power leaving unchanged other ratios.

Focusing on Italian firms, Pederzoli and Torricelli (2010) adopt a logit model to predict the default probability for a specific region, i.e. Emilia Romagna, based on financial ratios. They find that the equity ratio, the EBIT over asset ratio, the long term liabilities over asset ratio and sales over asset ratio are sufficient to fit the default event in their sample. In a similar context, for a sample of 232 Italian SMEs, Dainelli et al. (2013) develop a logit model for a one-year estimation of the probability of default. In addition to standard financial indicators (profitability, solvency and liquidity) they include credit relationship quality indicators. They find that both profitability and credit relationship quality are important determinants of the probability of default. Ciampi and Gordini (2013a) provide a methodological contribution and focus on a large sample of 7,000 Italian small enterprises (those with a turnover less than 1.8 million euro). They compare results obtained with different techniques and find that neural network analysis make a better contribution of the small enterprises credit risk evaluation when compared to traditional techniques, i.e. MDA and LR. Their prediction model is based on an appropriate set of financial and economic indicators.

Although these analyses differ among them, they converge toward the idea that five groups of financial indicators are most important in predicting default probability of SMEs: leverage, liquidity, profitability, coverage and activity ratios.

3. Dataset description and summary statistics[5]

Most of the existing models on the probability of default use public market data and may be used consequently for large (listed) firms. On the other hand, although SMEs constitute a high share of banks loans portfolio, data on them are not so easily accessible (Rikkers and Thibeault 2009).

In this paper, the sample used for the empirical analysis includes privately-held Italian SMEs provided by the *Centrale Rischi Finanizari* (CRIF), a credit rating agency issuing ratings on Italian firms, instrumental for banks and other financial institutions to evaluate their portfolios.

Building an appropriate database for our empirical analysis, it dealt with a number of issues. The first one has been on the definition of *SMEs*. To this end, we follow the Basel

[5] The database and the methodology used in this paper partially mirrors those adopted in a previous contribution by Muscettola and Pietrovito (2012).

II rules and the definition provided by the European Union (Commission recommendation 967280/EC) that counts the number of both employees and sales.

Firms with less than 250 employees and sales lower than 50 million euros are considered SMEs.[6] These requisites have been recently confirmed also by Basel III. The second issue concerns the default definition (Gai 2008).[7] As highlighted in Section 1, we adopt a broad definition of default reflecting that provided by the Basel II committee. Specifically, according to Basel II the default includes "non-performing exposures", i.e., interest or principal payments more than 90 days past-due. In the definition of default firms we include those showing 180 days past-due loans instead of 90-days, following the past-due trigger provided by the Italian supervisor. Moreover, this definition includes suffering, bankruptcy and ascertained losses. Regarding the economic activity, our sample includes firms operating in the manufacturing, commercial and service sectors. Data acquired includes balance sheets and income statements for four consecutive years, from 2006 to 2009.

After dealing with missing data and influencing observations, it has been created a sample of 9,208 small and medium firms. The analysis, therefore, is carried out on balance sheets and income statements of 2006 (2007) in order to analyze the characteristics of firms affecting their probability of default after three years, i.e., in 2009 (2010). By applying the default definition provided, the work focuses on two groups of firms: *default firms*, those showing suffering, 180 days past-due loans, bankruptcy and ascertained losses in either 2009 or 2010; and *non-default firms*, those not showing suffering, 180 days past-due loans, bankruptcy and ascertained losses until 2010. The composition of the samples of default and non-default firms is provided in Table 1. The firms included in the sample operate in different geographical areas, in different business sectors and differ in size.

The default group includes 322 firms in default in 2009 or 2010 and represents 3.5% of the total sample, while the non-default group consists of 8,886 companies representing 96.5% of the total. Concerning the geographical area where firms are located, the sample includes 6,330 firms in the North, 1,667 in the Center and 1,211 in the South of Italy. The distribution of default firms among different geographical areas mirrors the composition of the whole sample. Most part of default firms, at least in absolute terms, is concentrated

[6] In our database, the definition of SMEs refers to 2009.

[7] The literature (see, for instance, Gai 2008) provides several definitions of the default event. This choice is not neutral since it might influence the results of the empirical analysis, other than the number of firms classified as *default*. These definitions include legal "bankruptcy" or "liquidation" that are more restrictive in that the moment of legal failure does not reflect the real failure event.

in the North (218), while the numbers of default firms in the Centre and in the South are equal to 66 and 38, respectively.

An analysis of the distribution of firms across sectors of economic activity shows that approximately 44% of them operates in the manufacturing sector (4,074), 38% in the commercial sector (3,503 firms), while the remaining 18% in the service sector (1,631).

Table 1 also shows the composition of the sample depending on firm size. To this end, we choose the two parameters adopted by the Basel II committee to define SMEs: sales and number of employees. The table reveals that the default event may more likely occur in larger firms. Finally, regarding the distribution of the default events depending firm years of activity, it is biased towards younger firms: more than 50% of default events concerns firms with less than 15 years of activity.

In sum, the distribution of default events shows a high concentration in the group of large and young firms. Most notably, the distribution of the two groups of firms show a prevalence of default firms in the Central Italy and in the manufacturing sector.

The initial set of economic and financial indicators to predict the default event is selected on the basis of three criteria: their frequency in the literature (Barontini 2000; Altman and Sabato 2007; Muscettola and Naccarato 2012; Ciampi and Gordini 2013); their ability to describe essential aspects of company's economic and financial profile, namely, profitability, efficiency, leverage and liquidity; and most notably, the representation of the financial structure of firms considering indicators measuring different financial sources (bank loans, commercial debts and equity), their cost (including interest expenses) and their maturity.

Table 2 shows the descriptive statistics of these indicators, distinguishing default and non-default firms and balance sheets in 2006 and 2007. Default firms show on average a low profitability than non-default firms: for instance in 2006, Return on Equity shows and average value for default firms equal to 0.848 versus 7.423 for non-default firms and return on investment is equal to 5.631 for default firms and 7.438 for non-default firms. The same applies in 2007. Regarding efficiency indicators, default firms show a high incidence of amortization and fixed assets (both in 2006 and in 2007) and a higher credits and debts turnover (Cost of sales/Accounts payable and Sales/Accounts receivable) in 2006 for default firms in 2009 and a lower turnover in 2007 for default firms in 2010 with respect to non-default firms in both years. Most notably, the financial structure indicators show that default firms have a higher incidence of liabilities both on sales and on total assets and, concerning debt maturity, a higher incidence of short- and long-term debts. This fact occurs both in 2006 and 2007. Eventually, in terms of liquidity, default firms show a lower liquidity index, Cash and cash equivalents/Total assets, (4.008 versus 8.380) and a lower acid test (77.995 versus 102.644) in 2006 with respect to non-default firms.

4. Methodology and results

This section describes the steps we follow in the empirical analysis in order to provide a model of default prediction for our sample of 9,208 firms. After constructing 48 financial and economic ratios described in the previous section, we reduce this set via *the factor analysis*. In the second step, we estimate a *stepwise* logistic regression to select those factors that, on the whole, are significant in predicting the probability of default in a three years window. Finally, we validate the results by estimating a *bootstrap* logistic regression and we provide robustness checks by including in the logistic regression the original set of indicators, instead of factors.

4.1 Factor analysis

In this paper, the factor analysis (Spearman 1904; Hamilton 2009; De Laurentis et al. 2010), conducted with the principal-component factor method to analyze the correlation matrix, aims to simplify and reduce the initial set of 48 indicators through the extraction of a new set of unobserved variables (called "factors") statistically significant. We believe that the use of factor analysis in this paper is appropriate for three reasons: (*i*) financial and economic indicators are constructed from a large number of components and therefore they are likely to be closely related; (*ii*) the selection of which balance sheet and income statement items or their ratios might be difficult to identify; and (*iii*) multivariate statistics techniques can be used to isolate the essential "factors" from balance sheets and income statement items that are likely to be driving firm default.

The factor analysis is conducted on the original standardized indicators, i.e. with mean zero and variance equal to one. Table 3 reports factors extracted and their corresponding eigenvalues, i.e. the percentage of the total variance explained by each factor. The first factor is identified so as to explain the highest percentage of the total variance of the original variables, the second factor so as to explain a lower percentage, and so on. Factors with an eigenvalue greater than 1.1 are retained, so that they synthesize at least two original indicators. Factor F1 explains 23.2% of the total variance and factor F2, as expected, explains a lower percentage equal to 16.7%. The remaining factors explain lower percentage of the total variance. Overall, the 11 factors considered explain 83% of the total variability.

In order to properly interpret each factor and eventually give it a label, in the second step we estimate the factor loadings, i.e. the weight of each original variable on the factors extracted and the correlations between the factors extracted and the initial indicators. We extract the factor loadings rotated by using the Varimax method and the Kaiser Normalization (Kaiser 1958).[8] The matrix containing the rotated factor loadings for each original indicator is reported in Table 4. The identity and meaning of each extracted factor reflects those original ratios that show the highest impact (i.e., highest factor loading) on the factor. Variables with large loadings for the same factors are grouped and small factor loadings (below 0.5) are omitted.

Nevertheless, estimated factor represents a specific economic or financial characteristic for firms (Canbas et al. 2005).

For instance, the first factor *capitalization* reflects the ability of a firm to use stockholers' equity to finance investments. The most influential indicators are indeed those measuring the incidence of total equity on total assets and the incidence of total liabilities on total assets. This factor includes nine indicators. Those showing a negative sign indicate that a higher score of these ratios decreases the value of the factor; on the contrary, indicators with a positive sign have a positive impact on the factor.

The second factor, *efficiency*, includes nine indicators measuring both the profitability and the level of self-financing of the firm. All indicators show a positive sign meaning that an increase in both the profitability and the self-financing determines an increase of the factor.

The factor called *fixed assets* contains seven original indicators: those with the highest factor loadings refer indeed to the fixed capital of the firm. A higher incidence of fixed assets or amortization increases the value of this factor.

The forth factor *commercial working capital* is highly and positively influenced by two indicators of the inventories. Moreover, this factor includes also, with a negative sign, the inventory turnover and the account receivables.

The fifth factor, *interest expense and bank debt*, contains four original ratios that summarize the degree of bank debt and impact of financial charges. A higher value of these indicators translates in a higher value of this factor.

The sixth factor *industrial profitability* includes four indicators that indicate the ability of the industrial management to generate earnings.

[8] This method maximizes the variance of the squared loadings of each factor on all variables, and allows the differentiation between the original variables and the factors extracted. In other words, the Varimax rotation method minimizes the number of original variables that have a high weight on a specific factor and produces orthogonal factors, i.e. not related to each other.

The factor called *debt* includes both short term and long term debts: the first indicator is negative and the second positive.

Moreover, the factor *fixed assets coverage* includes two indicators with a positive sign: the long-term debt on fixed assets and equity on fixed assets.

The ninth factor *liquidity* includes the incidence of cash and cash equivalents on total assets and on sales, respectively. A higher loading of liquidity impacts positively on the factor.

The factor *rotation of receivables and payables* includes two ratios: cost of sales on accounts payable and sales on accounts receivable.

Finally, the last factor *incidence of costs* includes the incidence on the operating margin of wages and interest expenses.

In synthesis, the new variables created through the factor analysis summarize the essential characteristics of firms linked to the composition of sources of funding, the profitability, the efficiency, the relationship between assets and liabilities and the ability to generate liquidity. We use these new variables in the subsequent steps of the analysis in order to generate a model of default prediction. To this end, we predict the factor scores, i.e. the values of the new variables for each firm in the sample. Table 5 shows the descriptive statistics of the extracted factors for the two samples of firms and the distance between the means of the two groups of companies. These results are obtained from the analysis of the standardized mean. This table show that on average, default firm show higher values of the factors summarizing fixed assets, commercial working capital, industrial profitability, interest expense and bank debt, debt and incidence of costs. On the contrary, factors with a "positive" meaning, i.e. capitalization, efficiency, fixed assets coverage, liquidity and rotation of receivables and payables, show a higher value for non-default firms.

4.2 Logistic regression

At this stage, we use the logistic regression model to estimate the probability of default in a three years window of the sample of firms by using the factors extracted in the previous step as explanatory variables.

The logistic model assumes a relationship between an independent variable (not observed) and a set of observable variables associated to the event. In our case, the event is the default probability of firms, while the observable variables include the factors extracted from the initial set of financial and economic indicators.

Denoting by p_i the default probability after three years for firm *i*-th, with x_m the set of *m* explanatory variables, with α the constant term and with β_m the coefficients of the explanatory variables, the logistic model expresses the probability as a linear function of the economic and financial indicators in the following way:

$$\log it(p_i) = \ln(\frac{p_i}{1-p_i}) = \alpha + \beta_1 x_1^i + \beta_2 x_2^i + \ldots + \beta_m x_m^i \tag{1}$$

The left-hand term is the logarithm of the odd ratio, i.e. the ratio between the probability of default and that of non-default. The model parameters (β_m) are estimated by the maximum likelihood function: the higher the value of the parameter estimated, the higher the contribution of the explanatory variable x_m in predicting the probability of default.

From (1) we can obtain the probability of the default event p_i as follows:

$$p_i = \frac{1}{1+e^{-(\alpha+\beta_1 x_1^i + \beta_2 x_2^i + \ldots + \beta_m x_m^i)}} \tag{2}$$

Assuming that $z = \alpha + \beta_1 x_1^i + \beta_2 x_2^i + \ldots + \beta_m x_m^i$, the logistic function is as follows:

$$p_i = f(z) = \frac{1}{1+e^{-z}} \tag{3}$$

where, *f*(*z*) ranges between 0 and 1 and represents the probability of default in a 3-years window.

Equation (1) is estimated using the stepwise procedure to identify the best combination of significant explanatory variables in a regression (Shin and Lee 2002; Shin et al. 2005) and to include them in a model. As suggested by De Laurentis et al. (2010), it is not convenient to use the stepwise regression on the entire set of indicators, as this would make it difficult to achieve the convergence towards a stable pattern of prediction. The advantages of the factor analysis, instead of using the full set of explanatory variables from the balance sheet and income statement, rely indeed on the ease of obtaining such a convergence. The stepwise procedure allows to select those factors that are *all* significant in predicting the probability of default. In particular, we adopt the *forward selection* method that starts from the model including only the intercept (null model) and selects iteratively significant variables. At the end of the selection process, the convergence towards a model in which all the explanatory variables are jointly significant is obtained. Starting from the eleven factors extracted through the factor analysis, eight factors with a significance level between 1% and 5% are selected with this methodology.[9]

[9] Another selection algorithm, the *backward selection* method starting from the full model (including all explanatory variables), iteratively eliminates the explanatory variables that are not significant. Furthermore, the application of the backward methodology provides the same results,

The results of the stepwise logistic regression are reported in Table 6, where β indicates the estimated coefficients of the logistic regression. The coefficients mean the relation between each explanatory variable and the dependent variable, i.e. the logarithm of the odd ratio. In particular, each coefficient represents the change in the logarithm of the odd ratio generated by a unit change of the explanatory variable (holding constant other explanatory variables).

As it may be inferred from the table, among the 11 candidate factors, the stepwise procedure selects seven that are jointly significant in explaining the probability of default. The estimation results show that all explanatory variables have the expected impact and are significant at the 1% level.

The higher impact is exerted by the factor *interest expense and bank debt* with a coefficient equal to 0.889. This means that the interest expenses and the debts contracted with banks are the most influencing observable characteristics positively affecting the probability a firm is classified as default in the three years window. In general, Italian firms rely on banks loans to finance their investment and this costly financial strategy increases the likelihood they will experience a default event. The same conclusions apply for the factor *debt* showing a positive (and even lower) impact (0.225) of short and long term debts on the default probability. Another interesting result related to the financial structure of the firm is the level of *capitalization*. As in Table 6, this factor exerts indeed a negative impact on the probability of default with a coefficient equal to (-0.660). In other terms, the more the firm relies on equity, the lower the probability of experiencing a past-due event in the next three years.

A partial effect of it is the firm capability to finance the firm's fixed assets by using consolidate liabilities and stockholders' equity. The factor *fixed assets coverage* shows indeed a negative coefficient (-0.379), meaning that the higher the matching between long-term sources and assets, the lower the probability of default.

The *liquidity* factor shows, as expected, a negative and significant coefficient (-0.330): the higher this factor the lower is the probability of default in a three years framework. Similarly, *efficiency* shows a negative (-0.314) impact on the probability of default. Intuitively, the higher is the ability of the firm to self-finance (i.e. to use internal funds) its assets and the higher is its profitability, the lower is the likelihood of experiencing a default.

---

both in terms of coefficients and significance, and in terms of the performance of the model. These results available on request.

Concerning indicators measuring the performance of the firm in terms of industrial profitability and in terms of commercial working capital, we can observe a positive coefficient on both factors (0.221 and 0.254, respectively). A consistent commercial working capital, associated to a low inventory turnover and high levels of trade receivables, contributes to increase the default probability.

Nevertheless, the results of the logistic model are interesting in the sense that they highlight the observable characteristics of SMEs Italian firms that banks should take into account in order to decide which of them are eligible to be financed. According to our predicting model, the composition of the financial structure of the firm in terms of financial sources, their cost and their maturity is a relevant indicator of Italian SMEs' default. The capital structure seems to have a primary role in default prediction. An improvement of the financial structure of the company, resulting from a reduction of onerous debts and interest expenses, as well as a higher attention to stockholders' equity and new options of funding (e.g. quasi-equity instruments), result in a reduction of the risk of default. These characteristics are found to be more important than economic/industrial dimensions, such as the industrial profitability.

By applying the estimated parameters in equation (3), we can predict the probability of default for all firms in our sample, in a time frame of three years.

In order to evaluate the prediction accuracy of the model it's common in the literature to compute the correctly classified observations. Therefore, we classify a firm as *default* whenever its estimated probability of default (p_i) is higher that a given threshold, while we classify it as *non-default* otherwise. Such a classification of firms will in general differ from the observed default and non-default status of the firms in the sample (Bottazzi et al. 2011). In particular, we evaluate Type I and Type II errors prediction, as reported at the bottom of Table 6. Type I errors refer to firms that are actually defaulting, but are classified as non-default firms. On the contrary, Type II errors refer to non-defaulting firms that are incorrectly classified by the model as default.

For verifying the model performance there is a selection of different cut-off values to classify each firm as default or non-default according to our model. In particular, 0.3%, 0.5% 0.10% and 0.15% values have been used, with the lower value reflecting the total default rate in our sample (322 default firms out of 9,208). As argued by Bottazzi et al. (2011), it is standard to prefer prediction models reducing Type I errors, i.e. models maximizing the percentage of correctly classified default. For a bank it is indeed much more costly failing to predict a default than classifying a non-default firm as default. Table 6 shows that Type I errors increase with increasing cut-off values, whereas Type II decrease. More particularly, the percentage of correct default is higher when using the 0.03 cut-off, i.e. when it is fixed at the sample default rate. In this case, 63 default firms (over 322) are incorrectly classified as non-default. On the other hand, 3,164 non-default firms (over 8,886) are incorrectly classified as default.

5. Robustness checks

5.1 Model validation

As already argued by De Laurentis et al. (2010), in order to validate the results obtained through a prediction model, two distinct samples are required: an analysis sample – for the estimation of the model – and a validation sample – for the validation of results (including data that do not enter in the analysis sample). Since our sample includes 322 cases of default firms, a distinction between analysis sample and validation sample is not appropriate. For this reason, we adopt the bootstrap method to estimate the logistic stepwise regression and to validate the results of the model described in Section 4.

This methodology consists in iterating *n* times parameters (odd-ratios, standard errors, *p*-values and confidence intervals) estimations in the predicting model on a balanced sample of default and non-default firms. Since there are 322 default firms, we decide to fix the sample of default firms and randomly extract, with replacement, as many observations from the sample of non-default firms, re-estimating 100 times all regression parameters after each extraction.

After running 100 times the regression on the set of extracted factors, we collect all the parameters and report their average results on Table 7. This table shows that the factor most frequently entering the model is that related to the financial structure of the firm (i.e. *interest expense and bank debt*), followed by *industrial profitability, incidence of costs, capitalization*, *commercial working capital*, *efficiency* and *fixed assets*.

Reassuringly, all results, both in terms of the impact on the probability of default and in terms of significance, confirm those obtained before.

5.2 The logistic regression on the original indicators

We have discussed so far the main results of our logistic empirical analysis adopting as independent variables the factors extracted through the factor analysis from a set of 48 indicators constructed from balance sheets and income statements. Our main conclusion relies on the relevance of financial variables, with respect to economic variables, in predicting the default probability of a sample of Italian SMEs. In particular, we have shown the capital structure and the incidence of interest expenses as the most important determinants, alongside the capitalization of the firm.

However, the factor extraction through the factor analysis might be questionable for several reasons. *First*, the meaning and the identity of each factor is opaque in some cases. Just think, for instance, to the factor called *capitalization* that also includes two indicators of liquidity (i.e. acid test and liquidity), alongside variables that reflects the composition of the financial sources of the firm. *Second*, and most notably, default prediction models should help banks to predict which lending candidates are eligible to receive loans, and which not. Therefore, banks should be provided with a set of observable characteristics of the firms to be evaluated.

That's the reason why in this section we discuss the results of the stepwise logistic regression actually conducted on the original sample of indicators, that are observable by banks, instead of on factors. Results obtained with the forward selection algorithm, by fixing 0.01 as the level of significance for addition of variables to the model and 0.05 as the level for their removal, are reported in Table 8.

These results are consistent with those discussed in Section 4. Moreover, they are more specific in that they allow to identify the balance sheet or income statement indicator predicting default. Interestingly, we find that four out of eight indicators refer to debts, especially interest expenses and bank loans. The highest incidence is indeed exerted by the return on debt (i.e. the ratio between interest expense and total liabilities) with a positive and significant coefficient (0.396) at the 1% level. Similarly, short-term debt on sales, bank debt and total liabilities on assets present positive and highly significant coefficients. This means that the higher the debt exposure of the firm and the higher the interest to be paid on these debts, the higher the probability of firm default in the three years window. These results, are consistent with previous estimates and also with the literature (see, for instance, Bottazzi et al. 2011).

On the other hand, other explanatory variables show a negative impact on the default probability. For instance, concerning the commercial activity of the firm the accounts receivable shows a coefficient equal to -0.022 meaning that an increase in credits, probably indicating a higher level of sales, makes the firm more able to repay loans in the three-years framework. Surprisingly, the coefficient of the ratio between inventories and account receivables is negative (-0.001): probably, this result is due to the fact that most part of the variability of this indicator is already accounted for by the accounts receivables. Finally, consistent with our expectation, both fixed capital and the level of liquid assets are significantly associated to a lower probability of default.

6. Concluding remarks

This paper provides an empirical investigation of the determinants of firm default for a sample of 9,208 Italian limited liabilities SMEs in a time frame of three years, over the period 2006-2010, that partially includes the effects of the recent financial crisis. Our findings shows the relevance of financial variables as determinants of firm default. In particular, the strengthening of the capital structure is a determining factor in driving the ability of Italian SMEs to operate successfully in the medium term. Scilicet, a weak financial structure, characterized by a high level of indebtedness and limited supply of capital, with a higher incidence of interest expenses, is a determinant of future default events. The excessive firm debt and the high financial charges generate indeed a significant increase in the financial risk of Italian SMEs in a three-years window. Companies with a largest debt exposure have been found difficult to face the crisis and are more vulnerable from the intensification of the general market conditions. These results are robust to different empirical techniques.

In general, the micro and small firms have suffered a lot because their low capitalization levels makes more difficult to respond to the increasing globalization of markets and to the new credit criteria introduced by banks (Varaldo and Lamberti 2009). At the same time, the excessive bias towards debt makes the firm vulnerable to short-term factors that affect the stability of earnings and require constant readjustment interventions to achieve sustainable levels of leverage. In firms that rely on leverage, the large stock of debt and high borrowing costs raise the financial risk and weaken the ability to gather financial resources to support new investments. As a consequence, the growth patch decreases and the exposure to fluctuations in the credit market increases.

The policy implications of these results act on two sides. On one hand, bank are required to pay attention to the capital structure of SMEs and in particular to the composition of financial sources and the cost of debts, more than to economic variables. On the other side, it is a priority to intervene on the financial policy of SMEs. Companies should adopt more virtuous financial behaviour in which the target levels of debt are more consistent with the objective of long-term sustainability of their economic and financial performance. The need for more structured corporate finance is urgent. The current composition of financial sources (equity, bank debt, commercial debt) is no longer in line with the financial needs of businesses. New financial sources and new approaches to cover the financial needs must be enabled to get stronger (and dynamic) capital structure of enterprises and to promote effective access to the capital market (Maino and Modina 2012).[10]

The implications of this paper are especially true after the global financial crisis when it is difficult to think to the further raising of the bank debt stock considering the existing volumes and the downsizing of the role of bank credit in place. In a scenario where a good business strategy couldn't be accompanied by a weak financial policy, new funding opportunities, alternatives to banks, might improve the robustness of unlisted small and medium-sized firms and to mitigate the risk that a default occurs.

[10] In this context, the Italian legislator has introduced important innovations in terms of new financial instruments for enterprises. The law n. 134 passed by Italian Government on June 2012 has initiated the construction of a regulatory framework that allows the issuance of debt securities, debt securities (bonds or mini-bonds) and commercial paper by non-listed companies, which are unrated and in seeking funding and liquidity. This regulatory action can play a crucial role to strengthen the solidity of Italian companies because it moves towards a greater diversification of financial resources.

Acknowledgements

This paper is part of the research project "*Il rating tra impresa, banche e territorio*", financially supported by the Consorzio Universitario di Economia Industriale e Manageriale (CUEIM). First of all, we are extremely grateful to Renato Maino for his valuable supervision and useful comments on a previous version of this paper. We would also like to thank Marco Muscettola for his support in dataset preparation.

References

Altman E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609.

Altman E.I., Marco G., & Varetto F. (1994). Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, 18(3), 505-552.

Altman E. I., & Sabato G. (2005). Effects of the new Basel Capital Accord on bank capital requirements for SMEs. *Journal of Financial Services Research*, 28(1-3), 15-42.

Altman E. I., & Sabato G. (2007). Modeling credit risk for SMEs: evidence from the U.S. market. *ABACUS* 43(3), 332-357.

Barontini R. (2000). La valutazione del rischio di credito. Il Mulino.

Basel Committee on Banking Supervision (1999), A New Capital Adequacy Framework, June.

Basel Committee on Banking Supervision (2004), International Convergence of Capital Measurement and Capital Standards. A Revised Framework, June.

Beaver W.H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 71-111.

Becchetti L., & Sierra J. (2003). Bankruptcy risk and productive efficiency in manufacturing firms. *Journal of Banking and Finance*, 27(11), 2099-2120.

Beck, T., & Demirgüc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance*, 30(11), 2931-2943.

Bottazzi G., Grazzi M., Secchi A., & Tamagni F. (2011). Financial and economic determinants of firm default. *Journal of Evolutionary Economics*, 21(3), 373-406.

Canbas S., Cabuk A., & Kilic S.B. (2005). Prediction of commercial bank failure via multivariate statistical analysis of financial structures: the Turkish case. *European Journal of Operational Research*, 166, 528-546.

Ciampi F., & Gordini N. (2013a). Small enterprise default prediction modeling through artificial neural networks: An empirical analysis of Italian small enterprises. *Journal of Small Business Management*, 51(1), 23-45.

Ciampi F., & Gordini N. (2013b). Relazione impresa-territorio e modelli predittivi del default d'impresa. Primi risultati di un'analisi statistica sulle piccole imprese italiane. *Sinergie rivista di studi e ricerche*, 90, 51-76.

Dainelli F., Giunta F., & Cipollini F. (2013). Determinants of SME credit worthiness under Basel rules: the value of credit history information. *PSL Quarterly Review*, 66(264).

Demirgüç-Kunt A., Detragiache E., & Tressel T. (2008). Banking on the principles: Compliance with Basel Core Principles and bank soundness, *Journal of Financial Intermediation*, 17(4), 511-542.

De Laurentis G., & Maino R. (2009). *I rating a base statistica*. Bancaria Editrice.

De Laurentis G., Maino R., & Molteni L. (2010). *Developing, validating and using internal ratings. Methodologies and case studies*. Wiley.

Eisenbeis R. (1977). Pitfalls in the application of discriminant analysis in business. *The Journal of Finance*, 32(3), 875-900.

Foglia A., Laviola S., & Marullo Reedtz P. (1998). Multiple banking relationships and the fragility of corporate borrowers. *Journal of Banking and Finance*, 22(20), 1441-1456.

Gai L. (2008). Il rating delle PMI: un approccio metodologico per banche, confidi e intermediari finanziari. FrancoAngeli Editore.

Giannetti M. (2003). Do better institutions mitigate agency problems? Evidence from corporate finance choices. *Journal of Financial and Quantitative Analysis*, 38(1), 185-212.

Grice J. S., & Ingram R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research*, 54(1), 53-61.

Hamilton L. C. 2009. Statistics with Stata (Updated for Version 10). Belmont, CA: Brooks/Cole.

Härdle W., Lee Y.J., Schäfer D., & Yeh Y.R. (2009). Variable selection and oversampling in the use of smooth support vector machines for predicting the default risk of companies. *Journal of Forecasting*, 28(6), 512-534.

Jarrow R.A., & Turnbull S.M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, 50(1), 53-85.

Kaiser H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187–200.

Laviola S., & Trapanese M. (1997). Previsione delle insolvenze e qualità del credito bancario: Un'analisi statistica. Temi di discussione del Servizio Studi n. 318 Banca d'Italia, Roma.

Maino R., & Modina M. (2012). L'impatto dei sistemi di rating interno sulla politica di finanziamento delle imprese. *Banca Impresa Società*, 1, 89-106.

Merton R.C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2), 449-470.

Mossman C.E., Bell G.G., Swartz L.M. & Turtle H. (1998). An empirical comparison of bankruptcy models. *The Financial Review*, 33(2), 35-54.

Muscettola M., & Naccarato, F. (2012). Probability of Default and Probability of Excellence, an Inverse Model of Rating-One More Tool to Overcome the Crisis: An Empirical Analysis. *Business Systems Review*, 2(2), 71-93.

Muscettola M., & Pietrovito, F. (2012). La rilevanza delle variabili finanziarie nel rating: i risultati di un'analisi empirica sulle PMI italiane. *Sinergie Rapporti di ricerca*, 36, 325-347.

OECD (2005). SME and entrepreneurship outlook. http://www.oecd.org/industry/smes/oecdsmeandentrepreneurshipoutlook-2005edition.htm.

Ohlson J.A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.

Pederzoli C., & Torricelli, C. (2010). A parsimonious default prediction model for Italian SMEs. *Banks and Bank Systems*, 5(4), 28-32.

Pierri F., Burchi A., & Stanghellini E. (2011). La capacità predittiva degli indicatori di bilancio: una verifica sulle aziende umbre. Università di Perugia, Quaderni del Dipartimento di Economia, Finanza e statistica, n. 92/2011. http://www.ec.unipg.it/DEFS/uploads/qd_92web.pdf.

Piesse J., & Wood D. (1992). Issues in assessing MDA models of corporate failure: a research note. *The British Accounting Review*, 24(1), 33-42.

Platt H. D., & Platt M.B. (1990). Development of a class of stable predictive variables: the case of bankruptcy prediction. *Journal of Business Finance and Accounting*, 17(1), 31-51.

Ravi Kumar P., & Ravi V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. *European Journal of Operational Research*, 180(1), 1-28.

Rikkers F., & Thibeault A. (2009). A structural form default prediction model for SMEs, evidence from Dutch market. *Multinational Finance Journal*, 13(3/4), 229-264.

Shin K.S., Lee T.S., & Kim H.J. (2005). An application of support vector machines in bankruptcy prediction model *Expert Systems with Applications*, 28, 127-135.

Shin K.S., & Lee Y.J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23, 321.328.

Sironi A. (2003). Testing for Market Discipline in the European Banking Industry: Evidence from Subordinated Debt Issues. *Journal of Money, Credit and Banking*, 35(3), 443-72.

Spearman C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72-101.

Varaldo R., & Lamberti L. (2009). La grande crisi globale: una sfida per la politica industriale e per le imprese. *Economia e Politica Industriale*, 36, 9-19.

Varetto F. (1999). Metodi di previsione delle insolvenze: un'analisi comparata. *Il Rischio Creditizio. Misura e Controllo*, 178-301.

Table 1 – Sample composition[a]

	Default firms		Non-default firms	
	Number	*Percentage*	*Number*	*Percentage*
Total sample	322	3.50	8,886	96.50
Geographical area				
North West	115	3.24	3,437	96.76
North East	103	3.71	2,675	96.29
Centre	66	3.96	1,601	96.04
South	38	3.14	1,173	96.86
Sector of economic activity				
Manufacturing	162	3.98	3,912	96.02
Commerce	99	2.83	3,404	97.17
Services	61	3.74	1,570	96.26
Sales				
5 – 7.5 Million euros	45	2.64	1,660	97.36
7.5 – 12.5 Million euros	75	2.77	2,630	97.23
12.5 – 20 Million euros	96	3.65	2,532	96.35
20 – 50 Million euros	106	4.88	2,064	95.12
Number of employees				
Below 16	65	2.90	2180	97.10
17 – 35	60	2.49	2353	97.51
36 – 70	110	4.20	2506	95.80
Above than 71	87	4.50	1847	95.50
Years of activity				
Below 15	113	5.07	2,115	94.93
16 – 24	68	3.19	2,062	96.81
25 – 32	65	3.08	2,044	96.92
Above 33	76	2.77	2,665	97.23

[a] The first panel shows the number and percentage of firms belonging to the groups of default and non-default firms in 2009 and 2010. The second panel shows the number of firms belonging to the two groups, distinguished by geographical area. The third panel distinguishes firms by sector of economic activity. The fourth and fifth panels show the distinction of firms based on size: sales and number of employees. The last panel shows the number of firms distinguished by years of activity.

Table 2 – Initial set of indicators[a]

Panel A

		Default			Non-default		
	Indicator	*Mean*	*St. Dev.*	*Num.*	*Mean*	*St. Dev.*	*Num.*
Profitability	Cash flow/Invested capital	4.561	6.337	143	6.573	7.440	8,886
	Cash flow/Short-term debt	9.195	18.041	143	12.701	18.453	8,886
	Interest expense/Cash flow	129.362	211.960	143	53.361	155.421	8,886
	Net cash flow/Sales	4.090	5.898	143	4.255	5.597	8,886
	Return on Equity	0.848	33.238	143	7.423	36.584	8,886
	EBITDA/Sales	8.580	9.641	143	7.453	7.773	8,886
	Return on Sales	5.110	7.461	143	4.321	6.289	8,886
	Return on Investment	5.631	8.547	143	7.438	9.438	8,886
	Cost of sales /Sales	57.009	26.582	143	55.998	27.620	8,886
	Labor cost/EBITDA	1.429	3.690	143	2.111	4.189	8,886
	Interest expense/EBITDA	0.343	0.530	143	0.223	0.434	8,886
	Added value/Sales	21.621	15.187	143	20.252	15.331	8,886
	EBITDA/Invested capital	8.791	9.176	143	11.277	9.888	8,886
	EBIT/Total liabilities	6.970	11.599	143	10.553	14.336	8,886
	Labor costs/Sales	12.140	8.802	143	12.095	10.320	8,886
Efficiency	Amortization/Sales	3.307	4.116	143	2.962	3.943	8,886
	Cash and cash equivalents/Sales	4.442	7.770	143	6.214	9.881	8,886
	Fixed assets/Sales	34.014	50.237	143	27.707	60.713	8,886
	Net working capital/Sales	7.891	21.675	143	10.427	21.760	8,886
	Added value/Total assets	1.355	0.972	143	1.759	1.152	8,886
	Cost of sales/Accounts payable	10.650	52.861	143	7.845	37.006	8,886
	Sales/Accounts receivable	13.254	54.600	143	11.918	42.457	8,886
	Sales /Inventories	99.253	301.159	143	147.240	357.742	8,886
Leverage	Interest expense/Total liabilities	3.326	1.467	143	2.096	1.544	8,886
	Total liabilities/Sales	87.286	71.574	143	57.512	53.145	8,886
	Short-Term debt/ Total liabilities	82.481	16.404	143	87.931	15.984	8,886
	Total liabilities/Total assets	78.626	14.084	143	69.818	19.369	8,886
	Interest expense/Sales	2.655	1.922	143	1.247	1.502	8,886
	Short-Term debt/Sales	65.830	36.125	143	45.952	28.736	8,886
	Total equity/Sales	18.353	20.930	143	21.957	30.095	8,886
	Total equity /Total assets	17.480	13.338	143	24.601	18.298	8,886
	Long-term debt/Total assets	13.906	13.944	143	8.585	11.906	8,886
	Consolidate liabilities /Total assets	34.225	18.154	143	37.378	20.517	8,886
	Bank debt/Total assets	38.095	16.916	143	22.469	18.808	8,886
	Accounts payable/Total assets	30.463	16.462	143	35.056	19.615	8,886
	Total equity/Total assets	15.960	12.838	143	23.046	17.742	8,886
Liquidity	Sustainable assets /Total assets	26.970	16.961	143	23.356	19.431	8,886
	Fixed assets/Total assets	25.871	16.676	143	22.385	19.107	8,886
	Cash and cash equivalents/Total assets	4.008	5.790	143	8.380	11.381	8,886
	Accounts receivable/Total assets	42.846	18.123	143	47.808	21.797	8,886
	Inventories/Total assets	25.500	18.189	143	20.004	17.965	8,886
	Accounts receivable/Total assets	36.323	18.272	143	42.240	22.392	8,886
	Acid test	77.995	51.420	143	102.644	64.879	8,886
	Long-term debt/Fixed assets	295.088	760.401	143	459.710	934.798	8,886
	Current assets/Short-Term debt	122.674	56.002	143	141.332	73.240	8,886
	Total equity/Fixed assets	163.731	543.294	143	310.006	675.802	8,886
	Net working capital/Invested capital	8.155	19.794	143	16.932	26.277	8,886
	Inventories/ Accounts receivable	127.118	176.602	143	95.383	166.218	8,886

Panel B

	Indicator	Default Mean	Default St. Dev.	Default Num.	Non-default Mean	Non-default St. Dev.	Non-default Num.
Profitability	Cash flow/Invested capital	2.679	4.907	179	6.541	7.427	9,029
	Cash flow/Short-term debt	4.762	11.000	179	12.646	18.451	9,029
	Interest expense/cash flow	137.541	264.561	179	54.565	156.747	9,029
	Net cash flow/Sales	2.589	5.256	179	4.253	5.601	9,029
	Return on Equity	-7.227	47.959	179	7.319	36.541	9,029
	EBITDA /Sales	7.541	8.671	179	7.471	7.807	9,029
	Return on sales	3.861	6.826	179	4.334	6.310	9,029
	Return on investment	4.211	5.788	179	7.409	9.427	9,029
	Cost of sales/Sales	54.451	26.832	179	56.014	27.603	9,029
	Labor costs/EBITDA	2.361	5.564	179	2.100	4.182	9,029
	Interest expense/EBITDA	0.511	0.742	179	0.225	0.436	9,029
	Added value/Sales	20.915	15.192	179	20.273	15.329	9,029
	EBITDA/Invested capital	7.384	6.374	179	11.238	9.881	9,029
	EBIT/Total liabilities	4.645	7.754	179	10.496	14.303	9,029
	Labor costs/Sales	12.211	9.855	179	12.095	10.297	9,029
Efficiency	Amortization/Sales	3.431	4.251	179	2.968	3.946	9,029
	Cash and cash equivalents/Sales	3.831	7.830	179	6.186	9.854	9,029
	Fixed assets/Sales	44.269	80.639	179	27.807	60.564	9,029
	Net working capital/Sales	3.799	24.403	179	10.387	21.760	9,029
	Value of production/Total assets	1.364	0.947	179	1.753	1.151	9,029
	Cost of sales/Accounts payable	5.196	27.828	179	7.889	37.307	9,029
	Sales/Accounts receivable	6.701	27.708	179	11.939	42.673	9,029
	Sales /Inventories	114.100	330.728	179	146.480	356.952	9,029
Leverage	Interest expense/Total liabilities	3.769	1.688	179	2.116	1.551	9,029
	Total liabilities/Sales	89.870	74.471	179	57.984	53.610	9,029
	Short-Term debt/ Total liabilities	83.508	15.041	179	87.845	16.004	9,029
	Total liabilities/Total assets	80.980	13.792	179	69.957	19.328	9,029
	Interest expense/Sales	3.069	2.217	179	1.269	1.520	9,029
	Short-Term debt/Sales	67.598	34.940	179	46.266	28.971	9,029
	Total equity/Sales	20.335	34.637	179	21.900	29.975	9,029
	Total equity/Total assets	14.765	13.746	179	24.488	18.251	9,029
	Long-term debt/Total assets	13.175	11.964	179	8.669	11.959	9,029
	Consolidate liabilities/Total assets	30.659	17.049	179	37.328	20.485	9,029
	Bank debt/Total assets	40.129	17.787	179	22.717	18.880	9,029
	Accounts payable/Total assets	31.483	17.876	179	34.983	19.576	9,029
	Total equity/Total assets	13.579	11.570	179	22.933	17.696	9,029
Liquidity	Sustainable assets/Total assets	27.645	19.903	179	23.414	19.398	9,029
	Fixed assets/Total assets	25.558	19.110	179	22.440	19.075	9,029
	Cash and cash equivalents/Total assets	3.676	7.046	179	8.311	11.327	9,029
	Current assets/Total assets	43.151	19.247	179	47.729	21.751	9,029
	Inventories/Total assets	25.505	20.176	179	20.091	17.981	9,029
	Accounts receivable/Total assets	36.245	18.743	179	42.146	22.344	9,029
	Acid test	68.375	28.823	179	102.253	64.758	9,029
	Long-term debt/Fixed assets	245.093	435.702	179	457.102	932.482	9,029
	Current assets/Short-Term debt	112.715	44.898	179	141.037	73.034	9,029
	Total equity/Fixed assets	119.417	217.033	179	307.690	674.129	9,029
	Net working capital/Invested capital	3.307	18.050	179	16.793	26.209	9,029
	Inventories/Accounts receivable	122.295	174.044	179	95.885	166.424	9,029

[a] The table shows means and standard deviation of financial and economic indicators in 2006 (*Panel A*) for default and non-default firms in 2009 and 2007 (*Panel B*) for default and non-default firms in 2010. All indicators are in percentage.

Invested capital: Total assets minus financial fixed assets, cash and cash equivalent and unpaid receivables from shareholders.

Short-term debt: payables due within one year.

Long-term debt: any loan or debt obligation with a maturity of more than one year.

Current assets: cash or assets that could be converted into cash within one year.

Cash and cash equivalents: cash, checks and bank overdrafts.

EBITDA: a computation of a firm's earnings before interest, taxes, depreciation and amortization are deducted.

Table 3 – Eigenvalues of factors extracted[a]

Factor	*Eigenvalue*	*Proportion*	*Cumulative*
Factor1	11.137	0.232	0.232
Factor2	7.993	0.167	0.399
Factor3	3.761	0.078	0.477
Factor4	3.316	0.069	0.546
Factor5	2.700	0.056	0.602
Factor6	2.169	0.045	0.647
Factor7	1.795	0.037	0.685
Factor8	1.721	0.036	0.721
Factor9	1.575	0.033	0.754
Factor10	1.382	0.029	0.782
Factor11	1.204	0.025	0.807

[a] The first column of the table shows the factors extracted from the original standardized indicators through the factor analysis. The second column shows the eigenvalues of each factor, the third column the percentage of variance explained by each factor and the fourth column the cumulative variance.

Table 4 – Rotated factor loadings

	Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**	**Factor 7**	**Factor 8**	**Factor 9**	**Factor 10**	**Factor 11**
Variable	*Capitalization*	*Efficiency*	*Fixed assets*	*Commercial working capital*	*Interest expense and bank debt*	*Industrial profitability*	*Debt*	*Fixed assets coverage*	*Liquidity*	*Rotation of receivables and payables*	*Incidence of costs*
Cash flow/Invested capital		0.844									
Cash flow/Short-term debt		0.693									
Interest expense/Cash flow					0.516						
Net cash flow/Sales		0.687									
Return on Equity		0.518									
EBITDA/Sales		0.738									
Return on sales		0.844									
Return on investment		0.896									
Cost of sales/Sales						-0.652					
Labor costs/EBITDA											0.885
Interest expense/EBITDA											0.778
Added value/Sales						0.756					
EBITDA/Invested capital		0.906									
EBIT/Total liabilities		0.833									
Labor costs/Sales						0.808					
Amortization/Sales			0.751								
Cash and cash equivalents/Sales									0.822		
Fixed assets/Sales			0.874								
Net working capital/Sales	0.686										
Value of production/Total assets						-0.502					
Cost of sales/Accounts payable										0.710	
Sales/Accounts receivable										0.729	
Sales /Inventories				-0.557							
Interest expense/Total liabilities					0.714						
Total liabilities/Sales			0.667								
Short-Term debt/ Total liabilities							-0.898				
Total liabilities/Total assets	-0.863										
Interest expense/Sales					0.665						
Short-Term debt/Sales			0.535								
Total equity/Sales			0.668								
Total equity /Total assets	0.856										
Long-term debt/Total assets							0.920				
Consolidate liabilities/Total assets	0.803										
Bank debt/Total assets					0.640						
Accounts payable/Total assets	-0.531										

Total equity/Total assets	0.877										
Sustainable assets/Total assets			0.774								
Fixed assets/Total assets			0.767								
Cash and cash equivalents/Total assets									0.831		
Current assets/Total assets				-0.697							
Inventories/Total assets				0.872							
Accounts receivable/Total assets				-0.672							
Acid test	0.717										
Long-term debt/Fixed assets								0.892			
Current assets/Short-Term debt	0.812										
Total equity/Fixed assets								0.860			
Net working capital/Invested capital	0.734										
Inventories/Accounts receivable				0.800							

[a] The table shows the correlation coefficients between the original indicators and rotated factors (obtained with the Varimax method and the Kaiser Normalization). Variables with large loadings for the same factors are grouped and small factor loadings (below 0.5) are omitted

Table 5 – Descriptive statistics of factors extracted[a]

	Mean		
Factor	**Default**	**Non-default**	*Distance b*
Capitalization	-0.343	0.012	
Efficiency	-0.139	0.005	
Fixed assets	0.168	-0.006	
Commercial working capital	0.220	-0.008	
Interest expense and bank debt	0.971	-0.035	
Industrial profitability	0.189	-0.007	
Debt	0.157	-0.006	
Fixed assets coverage	-0.084	0.003	
Liquidity	-0.085	0.003	
Rotation of receivables and payables	-0.021	0.001	
Incidence of costs	0.061	-0.002	
Observations	322	8,886	

[a] This table shows the average values of factors extracted through the factor analysis for the sample o
and for the sample of 8,886 non-default firms, respectively. The statistics are computed on the standa

Table 6 – Stepwise logistic regression[a]

Factor	*β*	*S.E.*	*p-value*	*Exp(β)*
Capitalization	-0.660	0.107	0.000	0.517
Efficiency	-0.314	0.092	0.001	0.730
Commercial working capital	0.254	0.056	0.000	1.289
Interest expense and bank debt	0.889	0.052	0.000	2.433
Industrial profitability	0.221	0.060	0.000	1.248
Debt	0.225	0.049	0.000	1.252
Fixed assets coverage	-0.379	0.121	0.002	0.684
Liquidity	-0.330	0.088	0.000	0.719
Constant	-3.992	0.098	0.000	0.018
Cut-off	0.03	0.05	0.10	0.15
Type I Errors	63	116	213	262
Type II Errors	3,164	1,733	537	208
% Correct default	0.804	0.640	0.339	0.186
% Correct non-default	0.644	0.805	0.940	0.977

[a] This table shows the results of stepwise logistic regression conducted on the 11 extracted factors as determinants of the probability of default. The selection algorithm is the "forward" method. The table shows the coefficients (*β*) of factors with a significant level ranging between 1% and 5%. *S.E.* is the standard error of the estimated coefficients. *p-value* indicates the significance of the coefficient and *Exp*(*β*) indicates the linear prediction. Cut-off indicates the threshold used to define default and non-default firms in the sample by adopting the estimated model. Type I Errors refer to the number of firms that are actually defaulting, but are classified as non-default firms. Type II Errors refer to non-defaulting firms that are incorrectly classified as default. % Correct default indicates the percentage of firms that are correctly classified as default, over the total number of default (322). % Correct non-default indicates the percentage of firms that are correctly classified as non-default, over the total number of non-default (8,886).

Table 7 – Stepwise logistic regression with bootstrap[a]

Factor	*β*	*S.E.*	*p-value*	*Frequency*
Capitalization	0.303	0.457	0.008	4
Efficiency	1.265	0.238	0.006	2
Fixed assets	1.740	0.215	0.010	1
Commercial working capital	1.747	0.337	0.006	3
Interest expense and bank debt	2.580	0.271	0.002	74
Industrial profitability	2.316	0.290	0.005	5
Debt	2.020	0.242	0.004	1
Incidence of costs	1.450	0.237	0.002	5
Cut-off	0.03	0.05	0.1	0.15
Type I Errors	162	229	277	293
Type II Errors	1,409	687	271	131
% Correct default	0.497	0.289	0.140	0.090
% Correct non-default	0.841	0.923	0.970	0.985

[a] This table shows the results of the stepwise bootstrap logistic regression considering the 11 extracted factors. Balanced sample of 322 observations is used. The number of replications is 100. The selection algorithm is the "forward" method. The table shows the linear prediction *Exp*(*β*) for coefficients of factors with a significant level ranging between 1% and 5%. *S.E.* is the standard error of the estimated coefficients. *p-value* indicates the significance of the coefficient. *Frequency* indicates the number of times each factor is selected to be included in the model. Cut-off indicates the threshold used to define default and non-default firms in the sample by adopting the estimated model. Type I Errors refer to the number of firms that are actually defaulting, but are classified as non-default firms. Type II Errors refer to non-defaulting firms that are incorrectly classified as default. % Correct default indicates the percentage of firms that are correctly classified as default, over the total number of default (322). % Correct non-default indicates the percentage of firms that are correctly classified as non-default, over the total number of non-default (8,886).

Table 8 – Stepwise logistic regression on the sample of original indicator[a]

Indicator	*β*	*S.E.*	*p-value*	*Exp(β)*
Interest expense/Total liabilities	0.396	0.032	0.000	1.486
Short-term debt/Sales	0.015	0.002	0.000	1.015
Bank debt/Total assets	0.013	0.003	0.000	1.013
Total liabilities/Total assets	0.023	0.006	0.000	1.024
Accounts receivable/Total assets	-0.022	0.004	0.000	0.979
Fixed assets/Sales	-0.005	0.001	0.000	0.995
Cash and cash equivalents/Total assets	-0.038	0.011	0.001	0.962
Inventories/Accounts receivable	-0.001	0.000	0.009	0.999
Constant	-6.072	0.472	0.000	0.002
Cut-off	0.03	0.05	0.10	0.15
Type I Errors	60	106	202	254
Type II Errors	2,959	1,735	582	237
% Correct default	0.814	0.671	0.373	0.211
% Correct non-default	0.667	0.805	0.935	0.973

[a] This table shows the results of stepwise logistic regression conducted on the original set of indicators. The selection algorithm is the "forward" method. The table shows the coefficients (*β*) of factors with a significant level ranging between 1% and 5%. *S.E.* is the standard error of the estimated coefficients. *p-value* indicates the significance of the coefficient and *Exp*(*β*) indicates the linear prediction. Cut-off indicates the threshold used to define default and non-default firms in the sample by adopting the estimated model. Type I Errors refer to the number of firms that are actually defaulting, but are classified as non-default firms. Type II Errors refer to non-defaulting firms that are incorrectly classified as default. % Correct default indicates the percentage of firms that are correctly classified as default, over the total number of default (322). % Correct non-default indicates the percentage of firms that are correctly classified as non-default, over the total number of non-default (8,886).