The effect of size on the failure probabilities of SMEs: An empirical study on the US market using discrete hazard model

Izidin El Kalak[1,2]*
Hull University Business School
Email: i.el-kalak@2012.hull.ac.uk.

Robert Hudson
Hull University Business School
Email: robert.hudson@hull.ac.uk

Abstract

This paper investigates the extent to which the size affects the SMEs probabilities of bankruptcy. Using a dataset of (11,117) US non-financial firms, of which (465) filed for insolvency under chapters 7/11 between 1980 and 2013. We forecast the bankruptcy probabilities by developing four discrete-time duration-dependant hazard models for SMEs, Micro, Small, and Medium firms. A comparison of the default prediction models for medium firms and SMEs suggest that an almost identical set of explanatory variables affect the default probabilities leading us to believe that treating each of these groups separately has no material impact on the decision making process. However, comparisons between the micro and small firms with the SMEs firms strongly suggest that these categories need to be considered separately when modelling their credit risk.

JEL classification:G200

Keywords: SMEs; Bankruptcy; Size; Discrete-time Duration-dependant hazard model.

---

[1] Corresponding Author:robert.hudson@hull.ac.uk.
[2] Izidin El Kalak gratefully acknowledges financial support from the University of Hull.

1. Introduction:

Small and medium-sized enterprises (SMEs) are viewed as the backbone of the economy of many countries all over the world since they are the incubators of employment, growth, and innovation (Altman and Sabato, 2007). SMEs play a vital role in the US economy where statistics from the "US Small Business Administration[3]" show that small businesses make up 99.7% of US employer firms in 2011, and they accounted for 63% of the new jobs created between 1993 and 2013. These numbers emphasize the importance of SMEs as job creation engines; Furthermore, the Bureau of Labour Statistics[4] and a study by the economist intelligence unit in 2009 show that during the financial crises SMEs continued to hire employees and create new job opportunities (Economist intelligence Unit, 2009).

The introduction of the new Basel Capital accord and the global financial crises of 2007 opened the door for more in-depth and adequate research on failure[5] prediction models for all firms. However, the financial distress definition of Basel II, 90 days overdue on credit agreement payments, which is considered as the operational definition failed to distinguish between large and small firms which have different structure from credit risk point of view (Dietsch and Petey, 2004; and Altman and Sabato, 2007).

Credit risk modelling for large, listed firms is extensive and gravitates towards two approaches: The Altman (1968) approach which uses historical accounting data to predict bankruptcy; and the Merton (1974) approach which relies on securities market information.

More recently, banks and financial institutions started to realize the importance of distinguishing SMEs from large firms while modelling credit risk since they require specific risk management tools and methodologies to be developed for them (Altman et al., 2010). In line with this, Dietsch and Petey (2004) argue that German and French SMEs are riskier than large firms but have lower asset correlation with each other. Altman and Sabato (2007) provide a

[3] Small Business Administration known as "SBA" was created in 1953 as an independent agency of the federal government to aid, counsel, assist and protect the interests of small businesses in the US. For more details: http://www.sba.gov/

[4] Source: Bureau of Labour Statistics, BED. For the latest employment statistics, see Advocacy's quarterly reports, www.sba.gov/advocacy/10871.

[5] The terms failure, bankruptcy, default, and insolvency are used interchangeably in this paper.

distress prediction model specifically designed for the US SMEs sector based on a set of financial ratios derived from accounting information.

In recent years a new strand of literature has started to focus on the diversity within the SME category dividing the SMEs into micro, small, and medium sized firms. These categories are classified in terms of the firms' management style (Wager, 1998), access to finance (Beck et al., 2006), number of employees (Gupta et al., 2014) etc. A limited literature in this area has been devoted to studying the credit risk behaviour of these different categories (see for example, Gupta et al., 2014). In our study we will address this research gap by classifying SMEs into three distinct categories (micro, small, and medium) while developing a bankruptcy prediction model using a set of financial ratios. We will apply the discrete-time duration-dependant hazard rate modelling technique to develop separate bankruptcy prediction models for each of the three categories.

The main contribution of this paper is to investigate the extent to which the size affects the SMEs probabilities of bankruptcy by dividing our sample into three main size segments namely micro, small, and medium. In addition, we forecast the bankruptcy probabilities by developing discrete-time duration-dependant hazard models. Our paper is a continuation and improvement on three papers in the literature about SMEs failure: Altman and Sabato (2007), Holmes et al. (2010), and Gupta et al. (2014). We differ from Altman and Sabato's (2007) paper in two ways. Firstly, we classify SMEs into three categories (micro, small, and medium) while modelling for bankruptcy prediction. We try to capture any differences that exist between these categories and to what extent this might help lenders to further improve their credit models. Secondly, we utilize a more recent sample period (in and out of sample) which includes the recent financial crises in 2007, by doing this we assess the extent to which the financial crises affected the SMEs sector and the bankruptcy prediction model of SME firms. Holmes et al. (2010) study the survival of SMEs for the period from 1973 till 2001 and distinguish between micro firms and small and medium firms using hazard model methodology. They find that each segment is affected differently by firm-specific and macro-economic factors. However, the data used in their study differs significantly from our data, as they concentrate their sample on a specific geographical location within the UK (North-East England) and limited their sample only to a specific industrial segment which is the manufacturing sector and this sector represents only 12% of the UK firms. Moreover, they have not used any financial information in their analysis and cover a wide and back dated sampling period. We differ from the paper of Gupta et al.’s (2014) in several ways. First, we test the SMEs categories on a geographically different sample (US firms) and in doing so we emphasize the soundness and significance of distinguishing between the broad SMEs categories. Second, from a methodological point of view, while applying discreet hazard models, the estimation of baseline hazard should be done using time dummies (Beck et al., 1998) or some other functional form to model time (Jenkins, 2005). However, Gupta et al. (2014) have created the baseline hazard while including insolvency risk variable which distorts the idea of baseline hazard. Moreover, they utilize the ROC curve as their out of sample validation technique, however, this technique has been criticized by many scholars who argued it generates misleading results. In our study, we apply certain improvements to their paper by establishing a more precise baseline hazard function based on time dummies and apply an out of sample evaluation technique similar to the one used by Shumway (2001) which provides more accurate results.

Our analysis is carried out on a sample of (11,117) US non-financial firms of which (465) are defaulted firms, spanning a time period from 1980 till 2013. Our empirical findings show that significant differences exist between the bankruptcy attributes of micro and small firms on one hand and SMEs firms on the other. Therefore, separate treatment should be provided while modelling the credit risks of these categories. Moreover, we find similar results to those found by Gupta et al. (2014) in that the explanatory power of financial reports increases with the firm size. We find that medium and SMEs bankruptcy attributes have almost identical explanatory power leading us to believe that there is no material impact on the decision making process between these two groups unlike the micro and small SMEs. Finally, we provide an out of sample validation following the Shumway (2001) measure. Our out of sample results show good performance classifications for the four bankruptcy prediction models developed.

The remainder of the paper proceeds as follows. Section 2 provides an overview of the definition of SMEs, previous attempts to model failure probabilities for SMEs, and the studies conducted on micro, small, and medium-sized enterprises. Section 3 provides explanations about the source of the data used, the statistical methods utilized in this research, and the selection of covariates included in this study. Section 4 presents the key descriptive statistics for the covariates used and their correlation matrix, the univariate analysis applied, and the development of the discrete-time duration-dependant hazard models estimated for each of the SMEs segments. Finally, section 6 provides conclusions.

2. Literature Review

Research on small and medium-sized enterprises has gained a lot of attention and covered a wide range of issues in the previous decade. This section reviews the issues that are of particular relevance to this study which are the definition of SMEs, previous attempts to model failure probabilities for SMEs, and studies conducted on micro, small, and medium-sized enterprises.

2.1. SMEs Definition

To date, countries have failed to agree a general definition for small and medium sized enterprises. Therefore, each country defines their SMEs according to a particular set of firm-characteristics and quantitative variables. The most used variables in distinguishing small from large firms are the legal status, number of employees, independence, employment, industrial sector, asset size, and capital investment. The two main economic zones of interest to our study that provide detailed definitions for SMEs are the European Union and the US.

The Small Business Administration (SBA) is the main organisation that has been created by the US congress to deal with issues relating to SMEs. The SBA is also considered to be the major authority that defines SMEs in the US. A small business is defined in terms of the average number of employees and the average annual receipts. In addition, the SBA defines a number of other criteria to qualify as a small business: (i) is organised for profit (ii) has a place of business in the US (iii) Contributes to the US economy by paying taxes or using American products, materials, or labours (iv) independently owned and operated (v) does not exceed the numerical size standard for its industry[6]. In general, two widely used size standards have been established by the SBA, the maximum number of employees should be

[6] For detailed information about determining the business size, see http://www.sba.gov/content/determining-business-size

500 and the average annual receipts should be less than $7.5 million. However, there are a number of exceptions depending on the industry classification of the firm.

The 1996 law concerning the SMEs operating within the European Union Framework was updated in 2003 and provides a widely accepted definition for SMEs taking into account the new Basel rules. The law defines SMEs as firms having less than 250 employees with annual turnover of less than €50 million in sales.

2.2. Small and Medium-Sized Enterprises Failure

Measuring and tracking the probability of failure of small and medium-sized enterprises is a difficult task. This is due to the difficulties associated with locating and identifying these firms, in addition to determining the exact reasons for their failure (Altman et al., 2010). Despite the existence of these difficulties a considerable amount of research has been carried out to investigate the rates and causation of such failures (see for example, Watson and Everett (1993); Headd (2003); Carter and Auken (2006); Altman et al. (2010); and Gupta et al. (2014)).

The failure of new firms should not always be taken as economically inefficient, since it might enhance social welfare and reduce industry costs. In addition, according to Knott and Posen (2005) not all business failures are due to financial difficulties. Given this argument, before analysing business failure rates, it is essential to distinguish between firm failure and firm planned exit strategies where the business is actually healthy enough to continue operation (Headd (2003); and Bates (2005)). In line with this, Watson and Everett (1996) argue that some financially successful firms might decide to close for different reasons such as closing to limit losses, change of ownership, opportunity cost, switching costs, personal decisions etc.

Headd (2003) report that only one third of new businesses closure is due to circumstances that owner believed were due to a lack of success. Therefore in our study we have distinguished between failure due to financial constraints and closure due to strategic gains. Since the aim of this study is to develop a default prediction model for SMEs we separate the cases of failure from those of closure to improve the quality of the information available and the power of the prediction model and include only the firms that failed due to financial constraints.

The literature has further investigated the reasons behind business failures. Altman et al. (2010) mention two principle reasons for firms' closure which are lack of planning and insufficient capitalisation. Hutchinson and Xavier (2006) suggest that financial difficulties are the main factor for SMEs failure, while others such as Peacock (2000) report that poor managerial skills are behind these failures. Carter and Auken (2006) classify default factors into direct and indirect costs. They have suggested that the direct costs such as lack of knowledge, economic climate, and debt financing are the main reasons for firm failure, while indirect costs such as self-employment, personal collateral, self-esteem can play a secondary role.

In their paper Altman et al. (2010) suggest that different asset size segments lead to different SMEs insolvency risk behaviour. They find that the relationship between asset size and insolvency risk appears to be non-linear. They justify their argument that the lower the asset values the less likely the firm to be pursued by creditors for bankruptcy proceedings, since little opportunity remains for creditors to recover their debts. However, when the firm's assets value increases, insolvency proceedings become more attractive for creditors. Therefore, insolvency risk increases with increasing the asset size. However, after a certain threshold level this increase in bankruptcy risk starts to decline with additional increase in assets value. This finding is further supported in the literature that finds a non-monotone impact of size (for more details see, Brüderl et al. (1992); Falkenstein (2000); and Hamerle et al. (2006)). In line with the asset size argument, this paper tends to further classify SMEs into three distinctive categories (micro, small, and medium) believing that some factors leading to failure probability may vary across the three size categories. To our knowledge this is the first study conducted on the US market that develops a SMEs model for credit risk while distinguishing among micro, small, and medium firms.

2.3. Micro, Small, and Medium-Sized Enterprises

Micro, small and medium-sized enterprises (SMEs) are the engine of the economy. They are an essential source of jobs and create entrepreneurial spirit and innovation and are thus crucial for fostering competitiveness and employment. In 2005 a new definition from the EU came into force further classifying SMEs into three categories namely, micro, small, and medium enterprises. They defines a firm as 'micro' if it has less than 10 employees and an annual turnover of under €2 million; 'small' if it has less than 50 employees with an annual turnover of less than €10 million, and 'medium' if it has less than 250 employees with an annual turnover of less than €50 million. Since our paper aims to analyse the US market, we partially adopt these definitions and try to fit them within the SBA definition for SMEs in the US relying on the number of employees as the main factor of classification. Therefore, we will define a firm as 'micro' if it has less than 20 employees; 'small' if it has less than 100 employees; and 'medium' if it has less than 500 employees.

The empirical literature on SMEs has been extensively investigated especially after the new Basel Accord for bank capital adequacy (Basel II) (see for example Saurina and Trucharte (2004); Altman and Sabato (2005); and Berger (2006)). These studies covered a broad area of SMEs literature such as understanding the capital structure determinants of SMEs (Sogorb-Mira, 2005), investigating the key drivers of SME profitability and riskiness for US banks (Kolari and Shin, 2004) and the lending structure and strategies (Berger and Udell, 2004) etc.

Despite all these studies, a limited number of research studies have tried to further understand the sub-categories of SMEs and whether each category enjoys a unique set of characteristics. A study on the personnel management dimension within the SMEs conducted by Kotey and Slade (2005) show that differences exists between micro, small, and medium Australian firms. Their paper reports that the rate of adoption of formal human resources management practices increases with firm size. The results reported demonstrate a move toward division of labour, hierarchical structures, increased documentation, and more administrative processes as the number of employees increase. In addition, they advise taking into account the diversity of practices associated with various firm sizes and providing consultation and management training to SMEs personnel. Another study by De Mel et al. (2009) focused on the innovation dimension within the different categories of SMEs. They report that more than one quarter of microenterprises are found to be engaging in innovation, with marketing innovations the most common, and firm size is found to have a stronger positive effect, and competition a stronger negative effect, on process and organizational innovations than on product innovations.

Beck et al. (2005) investigate the effect of firm size on the extent to which the corruption of bank officials and financial and legal issue constrain a firm's growth. They found that the smaller the firm the more it is affected by these constrains. Besides differences in personnel management, innovation, and corruption, Beck et al. (2006) find that accessing finance also depends on the firm size, where they find that the larger the firm size the less access to finance is seen as a problem. They report that the probability a firm rates financing as a major obstacle toward its growth is 39% for small, 38% for medium, and 29% for large firms.

With regard to leverage decisions and capital structure, Ramalho and Da Silva (2009) conduct a study on Portuguese SMEs firms and show that different size structure (micro, small, medium, and large) significantly affect the determinants of leverage decisions. Research by Mateev et al. (2003) explores the capital structure choices for each of the SMEs categories. They find that medium-sized firms are mainly dependant on long term bank loans as their preferred method of external financing, while short-term loans and trade credits are the main source of external financing for both micro and small firms.

Recently, more attention has been given to the effect of SMEs categories on default probabilities and to what extent firm size matters in prediction of default. Empirical literature argues that the larger the firm is the more stable cash flow it holds and the more diversified it is (Gill et al., 2009) leading to a negative relationship between firm size and default probabilities (Pettit and Singer, 1985). A recent study by Gupta et al. (2013) investigates the financial and non-financial factors that influence failure within each of the SME categories (micro, small, and medium). Their findings provide strong evidence that the credit risk characteristics of firms within the broad SMEs segment do vary suggesting a separate treatment for each of the categories to get a better pricing of credit risk.

3. Empirical Analysis

This section provides detailed explanation about the source of the data used, the statistical methods utilized in this research, and the selection of covariates included in this study.

3.1. Data:

Our empirical analysis is performed using panel data from the Compustat database. The sample employs annual firm-level accounting data for (465) bankrupt and (11,117) non-bankrupt US small and medium-sized enterprises having less than 500 employees and an average annual receipts of less than $ 7.5 million, covering an analysis period from 1980 till 2013.

{Insert Table I here}

Furthermore, to validate the out-of-sample prediction performance of the models developed the entire study window is divided into two groups: the estimation period (1980-2008, 28 years) for the model building and the forecasting period (2009-2013, 5 years) for the out-of-sample forecasting performance test.

As discussed above, the SBA has established a widely used size standard to define SMEs of 500 employees and annual turnover receipts of $ 7.5 million for most industries. Moreover, the SMEs can be further classified into sub-samples of micro, small, and medium firms. The micro firms consists of less than 20 employees; firms are classified as "Small" if they have greater than or equal to 20 but less than 100; and "Medium" firms if they have greater than or equal to 100 and less than 500 employees. Further details regarding to the sub-samples are reported in table (I). It is important to mention that these definitions differ from the European Union ones which classify firms with only less than 250 employees as SMEs and which are used in different studies such as Altman et al. (2010) and Gupta (2014). Using our classifications, (213) failed micro firms are reported constituting around 46% of the total bankrupt SMEs sample compared to (115) failed firms for small SMEs and (137) failed firms for medium SMEs contributing 25% and 29% of the total bankrupt SMEs sample respectively.

{Insert table II}

In this study, we will consider firms to have failed only if they filed for legal bankruptcy proceedings (both Chapter 11 and 7) within the time period studied. Firms are classified as being legally bankrupt in Compustat database if the company has "TL" footnote on the status alert (Data item STALT) indicating that the firm is bankruptcy or in liquidation (e.g. Chapter 7/11). Furthermore, in line with other studies such as Altman et al. (2010) and Gupta et al. (2014), we exclude financial, insurance, and utility firms from our sample. The firms eliminated have industrial classification (SIC) codes from 6000 through 6999 for financial firms and 4900 through 4949 for regulated utilities. Finally, we will control for macroeconomic effects by including the changes in annual interest rates in the US throughout the period of our sample. This macroeconomic variable has been suggested by Hillegeist et al. (2001) as a control for macroeconomic conditions affecting the firm's default probabilities. In addition, we control for industry effects by classifying the firms into nine distinctive categories according to the SIC codes and including the variable as a factorial variable. Extreme outliers have been eliminated so that our models are not heavily influenced by them, we winsorised all our financial ratios between 5th and 95th percentiles. In addition, we have lagged all the covariates by one-time period so that all information is available in the beginning of the relevant time period.

{Insert table III}

3.2. Discrete-Time Duration-Dependant Hazard Model

3.2.1. The Hazard Model

In his seminal work Shumway (2001) argues that static models such as multiple discriminant analysis (MDA) and ordinary single-period logit techniques are inappropriate for default prediction due to the characteristics of bankruptcy data. The underlying characteristics for the majority of firms evolve over time but static models allow only for a single firm-year observation for each non-failed firm that is randomly drawn from the used data-set, while, for failed firms the firm-year observation immediately preceding the bankruptcy filing year is selected on a non-random basis leading to a possible sample selection bias (Hillegeist et al. 2004). Moreover, the single-period logit technique leads to understated values of standard errors (Beck et al. 1998), and fails to capture time-varying changes in the explanatory variable (Hillegeist et al. 2004). Therefore, researchers proposed new techniques to overcome the problems associated with static models. Hwang et al (2007) propose a robust semi-parametric logit model with smaller hold-out sample error rates. Whereas, Kukuk and Ronnberg (2013) suggest a mixed logit model which extends the normal logit model by allowing for varying stochastic parameters and non-linearity of covariates. Furthermore, Shumway (2001) suggests the utilization of hazard models in predicting bankruptcy probabilities where these models should be specified as duration dependant models with time-varying covariates. He highlights three reasons why the hazard model should be preferred over the static model: (i) the failure of the static logit to account for each firm's period at risk, (ii) the incorporation of time-varying explanatory variables, (iii) hazard models enjoy a higher predictive power in out-of-sample tests. Recent studies compare the Shumway model with other static models and show better forecasting performance of hazard models (see among others Chava and Jarrow (2004); and Bauer and Agarwal (2014)).

Furthermore, Hwang (2012) reports the superior performance of discrete-time duration-dependant hazard rate compared to the discrete-time hazard model without time-varying specification.

Nam et al. (2008) also argue that the discrete-time duration-dependant hazard model can be equivalent to a panel logistic model that incorporates macro-dependant base-line hazard.

The conditional probability of a discrete time hazard function (λ) for firm i to default in the time interval t, given it survives up to this time interval is as follows:

$$(\lambda\big(t\backslash X_{i,t}\big) = Pr(T = t\backslash T \geq t, X_{i,t})$$

T is discrete failure time; T = t states failure within the time interval t and $X_{i,t}$ is the value of the covariates of firm i up to time interval t, whereas the hazard model can be expressed in the following equation:

$$h(t\backslash X_{i,t}) = h(t\backslash 0)\,.\exp\{X_{i,t}^{`}\beta\}$$

Where, $h(t\backslash X_{i,t})$ is the individual hazard rate of firm i at time t and $X_{i,t}$ is the vector of covariates of each company i at time.

The discrete hazard technique fits well with the characteristics of the bankruptcy data utilized since it is consistent with the binary dependant variables and enjoys both time-series and cross-sectional characteristics. Furthermore, in line with the previous literature and to avoid the limitation of other statistical techniques we estimate our hazard models in a discrete-time framework with random effects (a_i) thus controlling for unobserved heterogeneity and shared frailty. The final equation used in this paper takes the following form, where a(t) is the time-varying covariate introduced to capture the baseline hazard rate and $P_{i,t}$ is the probability of experiencing the event by subject i and time t.

$$P_{i,t} = \frac{e^{\alpha(t)+\ \beta X_{i,t}}}{1+\ e^{\alpha(t)+\ \beta X_{i,t}}}$$

3.2.2. Specification of The Baseline Hazard Rate:

There are several ways to proxy the baseline hazard function a(t), when all the covariates are equal to zero, depending on the definition of the time-varying covariates that have functional relationships with survival times. The first method is the log (survival time) which has been applied by Shumway (2001) who used a time-invariant constant term, ln(Age). This is used for duration-independent models where the baseline hazard rate is assumed to be a constant term. In this case, the individual hazard rate, $\mathrm{h(t\backslash X_{i,t})}$ for firm i will be independent of the particular point of time or the survival period. The second method employs time dummies as a proxy for the baseline hazard rate. This method is utilized for duration-dependant models where the baseline hazard is assumed to be time-varying. Beck et al (1998) uses this method in their work, where the baseline hazard term, $\mathrm{k_t}$, is a dummy variable marking the length of the sequence of zeroes that precede the current observation. For example if the maximum survival time is sixty four years, then sixty three dummy variables are required for model estimation[7]. However, this method becomes more difficult if the maximum survival time in the dataset is very high as in the case of insolvency databases. Therefore, an alternative method to specify the baseline hazard rate is to use the piece-wise constant method. According to Jenkins (2005) this method splits the survival times into different time intervals that are each assumed to exhibit constant hazard rates. Overall, the choice of method depends on the shape of the hazard curve where frequent and continuous rises and falls suggest the use of fully non-parametric baseline hazard estimation.

Recently, some studies have moved away from baseline hazard estimation using time dummies by establishing other versions of baseline hazard that incorporates different types of variables. According to Nam et al (2008), indirect measures like time dummies are less effective in capturing time-varying macro dependences. Therefore, many researchers

[7] The model is run using sixty three years rather than sixty four dummies in order not to fall in the multicollinearity trap.

propose direct measures to estimate the baseline hazard rate. For example, Hillegeist et al (2001) propose the use of two direct measures; the rate of recent defaults and changes in interest rates. Nam et al (2008) use changes in interest rates and volatility of foreign exchange rates, whereas Altman et al (2010) and Gupta et al. (2014) construct industry "weight of evidence" variables.

3.2.3. Performance Evaluation

In order to examine the effectiveness of the models developed for the predition of SMEs bankruptcy we perform a bankruptcy out-of-sample prediction test similar to Shumway (2001). We specify our out-of-sample period to be from 2009 to 2013. Therefore, we re-calculate all the forecasting models for the period from 1980 till 2008 and then year by year we rank the firms into deciles based on their computed bankruptcy probabilities. The firms most likely to default in the subsequent year are placed into the first decile, the next most likely to default in the second decile, and so on. Subsequently, we report for each decile the percentage of firms that defaulted. The model is considered to enjoy better classification performance the higher the percentage of firms that experience default in the top deciles.

3.3. Selection of Covariates

A considerable number of ratios have been tested and used in the literature to predict SMEs default risk. Chen and Shimerda (1981) state that out of more than 100 financial ratios, almost 50% were found useful in at least one empirical study. This study focuses on the role of accounting ratios on the probability of SMEs failure. Therefore, the variables are selected from five broad categories that capture the firm's performance in the dimensions of profitability, leverage, activity, solvency, and liquidity. For each of these categories, we add a number of financial ratios that have previously been shown to be effective in predicting SMEs insolvency risk.

In order to select the most appropriate ratios for our final multivariate model, we apply two tests for each of the (20) financial ratios distributed over the five categories. Table (IV) presents the competing covariates that will be included in the univariate tests.

{Insert Table IV here}

The first step in choosing among these ratios is the implementation of a univariate regression analysis. This univariate test provides us with an initial understanding of the discriminatory power of the explanatory variables (Nam et al. (2008); Altman et al. (2010)). We keep all the ratios that show significant explanatory power and enjoy the expected sign relative to the dependant variable which is the probability of default. For the selected ratios we run a correlation test to identify any high correlations between these ratios. When ratios within each group exhibit high correlation, the covariates with lower chi-square values will be dropped from the final multivariate model since that indicates lower explanatory power for those ratios.

4. Results and Discussion

In this section we perform a univariate analysis of each individual covariate in our broad list of ratios followed by a correlation test. Furthermore, an analysis of key measures of descriptive statistics of the final selected explanatory variables is presented. Then we illustrate the process of developing our multivariate models for each SMEs category and for the SMEs as a whole. Thus allows us to compare and highlight the main differences between the models. Finally, we discuss the development of our out of sample classification performance for the models developed.

4.1. Univariate Analysis and Correlation Matrix

In this section univariate analysis is provided before proceeding to the development of the final multivariate models. Univariate analysis has been widely recommended and used in the literature to obtain an initial understanding of the discriminate power of the explanatory variables (Nam et al. (2008); Altman et al. (2010); and Gupta et al. (2014)). Usually, the standard approach in survival analysis is to obtain an insight about the shape of survival functions through the estimation of Kaplan-Meier survival curves for all categorical variables (Cleves et al., 2010). In addition, non-parametric tests such as log-rank and Wilcoxon-Breslow-Gehan tests are widely used to test the equality of survival functions for these categorical predictors (Cleves et al., 2010). However, the use of these tests may lead to biased discriminatory results if they have been applied on continuous predictors such as the case of our independent continuous variables[8]. So, to avoid any biased results univariate analysis will be conducted.

The results of the univariate regressions are reported in table (V).

{Insert Table V here}

To select the set of covariates that enter our multivariate model we choose those covariates that enjoy the expected sign while displaying significant discriminatory power when estimated using the discrete-hazard model for the different SMEs segments. An initial overview on table (V) indicates that within the profitability ratios all of the covariates, except for NISALE, RETA, and NITE have a significant discriminatory power and all those covariates show the expected sign compared to the dependant variable. However, among the leverage ratios, STDEBV and TDTA do not show the expected sign, at a significant level, relative to the probability of failure for all the three SMEs segments. Therefore, those two covariates are not considered during the next step. Regarding the remaining three ratio categories each of CG, WCSALE, CSIS, QCACL and CSIAT are not further considered in the correlation process because they do not provide enough statistical significance. Finally, after analysing the univariate regression for each covariates, the following covariates are tested to detect any

[8] See for example http://www.ats.ucla.edu/STAT/stata/seminars/stata_survival/default.htm. Also see Cleves *et al.* (2010) for a more thorough understanding.

multicollinearity, EBIDTAIE, EBIDTATA, NITA, XINTTA, CLTA, TCTA, TLTA, CETL, CASALE, TTA, WCTA, and CACL.

The correlation matrix is presented in table (VI) providing details about the collinearity level among the selected covariates. Out of the twelve covariates, the highest correlations can be found between EBIDTATA and NITA of about (0.9104), CACL and CETL (0.8314), CACL and WCTA (0.7789), TLTA and WCTA (-0.7577). A number of other covariates also have a substantial degree of correlation such as XINTTA and TLTA (0.6936) and CETL and CLTA (-0.6728). Some of the covariates have to be dropped from our final multivariate model due to the high correlations that exist between them. When two covariates are highly correlated with each other we keep the covariate that enjoys higher Wald chi-square value obtained from the univariate test table. Therefore, we determine seven covariates to enter the multivariate models namely, EBIDTAIE, NITA, TLTA, TCTA, CASALE, TTA, and WCTA.

{Insert Table VI here}

4.2. Descriptive Statistics

A discussion about the descriptive statistics of the covariates used in this study provides us with an initial understanding about any potential biases and variability that may arise among the variables in the multivariate models. In table (VII) we report the mean values and standard deviations for each of the three SMEs categories (micro, small, and medium) and for the whole sample separating the healthy and failed firms. A general overview of the descriptive analysis for the covariates selected shows initial evidence of differences among the variables in different SMEs categories which supports our argument that the factors influencing failure probability differ between each segment. For instance, the mean of EBIDTAIE differs among each category particularly between the SMEs which have (-4.518) mean value for failed firms and the medium failed firms with mean of (1.601) which might indicate that the profitability in medium failed firms is much higher than other groups. Surprisingly, the profitability of healthy micro and small SMEs have negative profitability ratios compared to healthy medium SMEs who enjoy a positive mean of (10.183).

In addition, the liquidity ratio WCTA among the micro and small failed firms provide negative results of (-0.006) and (-0.007) respectively, whereas it is positive among their peers in medium SMEs. This leads us to assume a liquidity problem among the micro and small failed firms compared to medium SMEs.

On the other hand, according to economic hypotheses and previous studies such as (Altman and Sabato (2007); Altman et al. (2010) etc.) we expect higher means in the failed group than for healthy group for the covariates that enjoy a positive relationship with the probability of failure. Not surprisingly, the means of the leverage ratios (TLTA) and (TCTA) for failed firms are higher than that for the firms in the healthy group among all the categories. Similarly lower means are expected for the covariates in the failed groups compared to those in the healthy groups when these covariates are negatively related to failure probability such as EBIDTAIE, NITA, CASALE, TLTA, and WCTA. Generally these expected relationships hold with the exception of that for TLTA.

{Insert Table VII here}

4.3. The Development of The Discrete-Time Duration-Dependant Hazard Models

In this section, we report on four hazard models that have been separately developed for SMEs, micro, small, and medium firms. The first step in this section is the detection of the baseline hazard rate which is the corner stone to further develop the discrete-time duration-dependant hazard models. This is followed by the development and discussion of the final multivariate models for each segment. The dependant variable for each model is a binary choice variable where (1) indicates bankruptcy and (0) indicates non-bankruptcy. The covariates selected to set up the multivariate models are chosen after consideration of their significance and correlation with other potential variables.

4.3.1. Determination of the Baseline Hazard Rate

The construction of the baseline hazard rate for these models can be done in different ways as explained in section (3.2.2.). However to chose between these methods the survival and hazard curves most be estimated and analysed. Figure (I) provides the estimated curves based on the Kaplan-Meier estimator for the four models separately. The survival probabilities for the whole SMEs model tend towards slightly above 0.50 as the firm age increases towards sixty. However, the survival probability for micro SMEs reduces to below 0.25 when the firms' age touches sixty years. Regarding the survival probability of small SMEs it moves to less than 0.50 when the firms' age approaches sixty years. In contrast to the small SMEs the survival probabilities of medium SMEs move in line with those of the SMEs taken as a whole to indicate survival probabilities of just above 0.50 at age 60. The different behaviours of the survival curves for each segment indicate that the survival attributes may be different for each size category. Even though the survival curves give us an initial understanding about the relationship between survival probabilities and the firms' age, it is important to plot the hazard curve for each model in order to decide the most appropriate method of calculating the baseline hazard. From figure (I) we can observe that different baseline hazard rate specifications are required for each model since each hazard curve exhibits a different functional relationship with firms' age. Moreover, since all the hazard curves show non-constant hazard rates for any defined age group a piecewise-constant method is inappropriate for this calculation, therefore we will use a fully non-parametric baseline hazard specification using age specific dummy variables to specify the baseline hazard rate. The minimum age of a firm in our sample is 1 while the maximum age is 64. Therefore, we generate 63 age specific dummies to represent all age categories.

{Insert Figure I here}

4.3.2. Discrete-time Duration-dependant Hazard Models

4.3.2.1. Hazard Model for all SMEs

The first model developed in this paper is the hazard model for all the SMEs in our sample which contain all the firms having less than 500 employees accounting for a total of 79,016 firm-year observations. In this model we have included all the covariates that are found to be significant during our univariate analysis. Table (VIII) provides the final results of the SMEs prediction model where it can be seen that all the covariates have coefficients of the expected sign. However, the NITA covariate fails to provide any significant discriminatory power in the multivariate setup.

4.3.2.2. Hazard Model for Micro Firms

This model has been estimated using the Micro SMEs' sample of firms that have less than 20 employees. Table (VIII) reports the final distress prediction model for Micro firms using the six selected covariates. The results in table (VIII) indicate that only two covariates show significant power in identifying the financial distress of micro SMEs namely TLTA and WCTA, whereas EBIDTAIE, NITA, TCTA, TTA, and CASALES exhibit insignificant power in the micro model. These findings are in line with the findings of Gupta et al. (2014) in the UK market that the explanatory power of financial reports increases with the size of the firm. We find that the larger the firm's size, the more similar are the results for the model to those for the SMEs model for all firms. In addition, after comparison between the small and medium models and the SMEs model, the empirical findings strongly suggest that the credit risk characteristics of micro SMEs differ from other SMEs and need to be considered separately when modelling their credit risks.

4.3.2.3. Hazard Model for Small Firms

This model has been estimated using the small SMEs' sample of firms which have has less than 100 and more than 20 employees. The results in table (VIII) indicate that four covariates in the model are insignificant in explaining the financial distress of Small firms namely EBIDTAIE, NITA, WCTA, and CASALES.

4.3.2.4. Hazard Model for Medium Firms

This model has been estimated using the medium firms' sample of firms which have less than 500 and more than 100 employees. The model for medium SMEs gives relatively similar results to the final model for SMEs showing highly significant covariates (except for NITA and TTA). After a comparison between the results of the two models, we can conclude that there are no strong reasons for creditors and decision makers to treat SMEs and medium firms separately.

4.4. Model Forecasting Accuracy

As mentioned in section (3) to test the effectiveness of the models developed in the predition of SMEs bankruptcy and their forecasting abilities table (IX) provides the classification performance measure for each of the prediction models developed. In terms of the models classification performance we find that all of the four models are able to capture more than 60% of the distress firms in the top three deciles which is considered to be a good percentage whereas the total number of the last five percentiles is less than 20%.

5. Conclusion

This paper investigates the extent to which the size of SMEs affects their probabilities of bankruptcy. To answer this question we classify SMEs into three size categories (micro, small, and medium) while modelling for bankruptcy prediction. We will try to capture any differences that exist among these categories and to what extent this might help lenders to improve their credit models.

We apply discrete-time duration-dependent hazard rate modelling techniques to develop separate bankruptcy prediction models for micro, small, and medium firms respectively, using a relatively large database of US firms. We compare their performance with the model developed for SMEs, as a whole including micro, small, and medium firms.

Our empirical analysis is performed using panel data available to us from the Compustat database. The sample employs annual firm-level accounting data for (465) bankrupt and (11,117) non-bankrupt US small and medium-sized enterprises having less than 500 employees and average annual receipts of less than $ 7.5 million, covering an analysis period from 1980 till 2013.

In order to test the effectiveness of the models developed in the perdition of SMEs bankruptcy and their forecasting abilities we perform a bankruptcy out-of-sample prediction test similar to Shumway (2001). We specify our out-of-sample period to be from 2009 to 2013. Therefore, we re-calculate all the forecasting models for the period from 1980 till 2008 and then year by year we rank the firms into deciles based on their computed bankruptcy probabilities. The firms most likely to default in the subsequent year are placed into the first decile, the next most likely to default in the second decile, and so on. Hence, the higher percentage of firms that experience default in the top deciles reflects a model with better classification performance. All the multivariate models developed exhibit strong classification performance capturing more than 60% of the distressed firms in the top three deciles which is considered to be a good percentage whereas the total number of the last five deciles is less than 20%.

A comparison of the default prediction models for medium SMEs and the whole SME sample suggest that an almost identical set of explanatory variables affect the default probabilities leading us to believe that there is no material impact on the decision making process by treating each of these groups separately. However, comparisons between the micro and small SMEs with the whole SME firms strongly suggest that they need to be considered separately when modelling credit risk for them. Based on our findings, we strongly advise lenders to provide a separate credit modelling assessment for micro and small SMEs since financial reports do not provide sufficient information about the likelihood of their default.

Reference

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

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

Altman E.I., Sabato, G. (2007). Modelling credit risk for SMEs: evidence from the US market. ABACUS 43, 332–357.

Altman E.I., Sabato, G., and Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk* 2:95–127

Bates, T. (2005). Analysis of young, small firms that have closed: delineating successful from unsuccessful closures. *Journal of Business Venture* 20:343–358

Bauer, J., and Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance* 40, 432–442.

Beaver, W. (1967). Financial ratios predictors of failure, empirical research in accounting: selected studies 1966. *Journal of Accounting Research* 4, 71–111.

Beck, N., Katz, J.N., and Tucker, R. (1998). Taking time seriously: time-series-cross-section analysis with a binary dependent variable. . *American Journal of Political Science* 42:1260–1288.

Beck, T., Demirguc-Kunt, A., and Maksimovic, V. (2005). Financial and legal constraints to growth: does firm size matter? *Journal of Finance* 60:137–177.

Beck T, Demirguc-Kunt, A., and Maksimovic, V. (2006). The determinants of financing obstacles. *Journal of International Money and Finance* 25:932–952.

Berger, A. (2006). "Potential competitive effects of Basel II on banks in SME credit markets in the united states. *Journal of Financial Services Research* Vol.28.

Berger, A., and Udell, G. (2004). A more complete conceptual framework about SME finance. World Bank conference on SME: Overcoming growth constraints, 14-15.

Brüderl, J., Preisendörfer, P., and Ziegler, R. (1992). Survival chances of newly founded business organizations. *American Sociological Review* 57, 227–242.

Carter, R., and Auken, H.V. (2006). Small firm bankruptcy. Journal of Small Business Management 44:493–512.

Chava, S. and Jarrow, R.A. (2004). Bankruptcy prediction with industry effects. Review of Finance 8 (4), 537–569.

Chen, K., and Shimerda, T. (1981). An Empirical Analysis of Useful Financial Ratios. Financial Management, Spring, 51-60

Craig, B. R., Jackson, W.E., and Thomson, J.B. (2004). On SBA-Guaranteed lending and economic growth. Working paper, Federal Reserve Bank of Cleveland.

De Mel, S., McKenzie, D., and Woodruff, C. (2009). Innovative firms or innovative owners? Determinants of innovation in micro, small, and medium enterprises. Policy Research Working Paper Series 4934. The World Bank.

Dietsch, M., and Petey, J. (2004). Should SME exposure be treated as retail or as corporate exposures? Comparative analysis of default probabilities and asset correlation in French and German SMEs. *Journal of Banking And Finance* 28, 773-788.

Economist Intelligence Unit (2009). Surviving the drought.

Falkenstein, E. (2000). Risk CalcTM for private companies: Moody's default model. Working Paper, Moody's Investors Service, New York.

Gill, A., Biger, N., Pai, C., and Bhutani, S. (2009). The determinants of capital structure in the service industry: evidence from United States. *The Open Business Journal* 2:48–53

Gupta, J., Gregoriou, A., Healy, J. (2014). Forecasting bankruptcy for SMEs using hazard function: to what extent does size matter? Review of Quantitative Finance and Accounting 1:25.

Hamerle, A., Liebig, T., and Scheule, H. (2006). Forecasting credit event frequency for loan portfolios. *The Journal of Risk* 9, 75–98.

Headd, B. (2003). Redefining business success: distinguishing between closure and failure. *Small Business Economics* 21, 51–61.

Hillegeist, S.A., Keating, E.K., Cram, D.P., and Lundstedt, K.G. (2001). Corporate bankruptcy: do debt covenant and disclosure quality measures provide information beyond options and other market variables? Working paper, Kellogg Graduate School of Management.

Hillegeist, S.A., Keating, E.K., Cram, D.P., and Lundstedt, K.G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies* 9 (1), 5–34.

Holmes, P., Hunt, A., and Stone, I. (2010). An analysis of new firm survival using a hazard function. *Journal of Applied Econometrics* 42:185–195.

Hutchinson, J., and Xavier, A. (2006). Comparing the impact of credit constraints on the growth of SMEs in a transition country with an established market economy. *Small Business Economics* 27:169–179.

Hwang, R.C. (2012). A varying-coefficient default model. *International Journal of Forecasting* 28:675–688.

Hwang, R., Cheng, K.F., Lee, J.C. (2007). A semiparametric method for predicting bankruptcy. *Journal of Forecasting* 26:317–342.

Jenkins, S.P. (2005). Survival analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Knott, A.M., and Posen, H.E. (2005). Is failure good? *Strategic Management Journal* 26:617–641.

Kolari, J.W., and Hwan, G.S. (2004). Assessing the profitability and riskiness of small business lenders in the US Banking industry. www.sba.gov.

Kotey, B., and Slade, P. (2005). Formal human resource management practices in small growing firms. *Journal of Small Business Management* 43:16–40.

Kukuk, M., Ronnberg, M. (2013). Corporate credit default models: a mixed logit approach. *Review of Quantitative Finance and Accounting* 40:1–17.

Mateev, M., Poutziouris, P., and Ivanov, K. (2013). On the determinants of SME capital structure in central and Eastern Europe: a dynamic panel analysis. *Research in International Business and Finance* 27, 28–51.

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

Nam, C.W., Kim, T.S., Park, N.J., and Lee, H.K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting* 27:493–506

Peacock, R. (2000). Failure and assistance of small firms. Available at http://www.sbeducation.info/downloads/sbfail.pdf

Pettit, R.R., and Singer, R.F. (1985). Small business finance: a research agenda. *Journal of Financial Management* 14:47–60.

Ramalho, J.J., Da Silva, J.V. (2009). A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms. *Journal of Quantitative Finance* 9:621–636.

Saurina, J., and Trucharte, C. (2004). The impact of Basel II on lending to small and medium-sized firms: A regulatory policy assessment based on Spanish credit register data. *Journal of Finance Service Research* 26, 121-144.

Shumway, T. (2001). Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business* 74:101–124.

Sogorb-Mira, F. (2005). How SME uniqueness affects capital structure: evidence from a 1994–1998 Spanish data panel. *Small Business Economics* 25:447–457.

Wager, T.H. (1998). Determinants of human resource management practices in small firms: some evidence from Atlantic Canada. *Journal of Small Business Management* 36:13–23.

Watson, J., Everett, J.E. (1996). Do small businesses have high failure rates? Evidence from Australian retailers. *Journal of Small Business Management* 34:45–62.

List of Tables and Figures

Table I: The composition of the sample of Bankrupt and Healthy firms

Number of firms by year					
Year	**Bankrupt firms**	**% of Total sample**	**Healthy firm**	**% of total sample**	**Total Sample**
1980	15	1.98	743	98.02	758
1981	4	0.98	403	99.02	407
1982	7	1.67	413	98.33	420
1983	15	3.59	403	96.41	418
1984	12	2.95	395	97.05	407
1985	13	2.73	463	97.27	476
1986	21	4.31	466	95.69	487
1987	18	4.74	362	95.26	380
1988	11	3.78	280	96.22	291
1989	20	7.69	240	92.31	260
1990	17	6.88	230	93.12	247
1991	24	6.82	328	93.18	352
1992	13	4.04	309	95.96	322
1993	23	5.42	401	94.58	424
1994	21	5.13	388	94.87	409
1995	19	3.91	467	96.09	486
1996	18	3.44	505	96.56	523
1997	23	6.12	353	93.88	376
1998	26	8.67	274	91.33	300
1999	17	2.96	558	97.04	575
2000	13	3.07	411	96.93	424
2001	19	7.79	225	92.21	244
2002	11	7.01	146	92.99	157
2003	14	6.97	187	93.03	201
2004	9	7.26	115	92.74	124
2005	13	8.50	140	91.50	153
2006	9	5.36	159	94.64	168
2007	4	2.31	169	97.69	173
2008	8	6.06	124	93.94	132
2009	10	6.37	147	93.63	157
2010	6	3.17	183	96.83	189
2011	4	1.95	201	98.05	205
2012	5	1.84	267	98.16	272
2013	3	1.50	197	98.50	200
Total	**465**	**4.18**	**10652**	**95.82**	**11117**

Table II: The distribution of US dataset across SMEs segments

Firm Category	Failed	Healthy	Total	Failed/Total %
SMEs	465	10652	11117	4.18
Micro	213	2638	2851	7.47
Small	115	3389	3504	3.28
Medium	137	4625	4762	2.88

The table shows the sub-classification of our database among micro, small, and medium firms, in addition to their default rate percentage

Table III: Industry code construction

IND Code	SIC code	Industry name	number of bankruptcies	% of Bankruptcies
1	<1000	Agriculture, Forestry and Fisheries	7	1.51%
2	1000 to less than 1500	Mineral Industries	36	7.74%
3	1500 to less than 1800	Construction Industries	14	3.01%
4	2000 to less than 4000	Manufacturing	186	40.00%
5	4000 to less than 4899	Transportation and Communications	36	7.74%
6	4950 to less than 5200	Wholesale Trade	28	6.02%
7	5200 to less than 6000	Retail Trade	40	8.60%
8	7000 to less than 8900	Service Industries	53	11.40%
9	9100 to less than 10000	Public Administration	65	13.98%
		Total # of Bankruptcies	**465**	100.00%

The above table gives the SIC codes a corresponding Industry codes along with the name of each industry. The last column gives the number of bankruptcies during the sample period 1980 - 2013 in each of these industries.

Table IV: Definition of variables

Code	**Definition**	**Compustat item code**
Profitability Ratios		
EBIDTAIE	Earnings before interest taxes depreciation and amortization/Interest expense	EBIDTA/XINT
EBIDTATA	Earnings before interest taxes depreciation and amortization/Total Assets	EBIDTA/AT
NISALE	Net income to net sales	NI/SALE
RETA	Retained earnings to Total assets	RE/AT
NITA	Net income to Total Assets	NI/AT
NITE	Net income to total equity	NI/TE
Leverage		
XINTTA	Financial Expenses/Total Assets	XINT/AT
CLTA	Total current liabilities/Total assets	LCT/AT
TCTA	Trade Creditors/Total Assets	AP/AT
TLTA	Total Liabilities to Total Assets	AT/AT
STDEBV	short term debt to equity book value	DLC/SEQ
TDTA	Total debt to total assets	DT/AT
Activity		
CETL	Capital employed/Total liabilities	(AT - LCT)/LT
TTA	Taxes/Total Assets	TXT/AT
CG	Capital Growth; Capital/Capital[_n-1]	((AT-LCT)/(AT[_n-1]-LCT[_n-1]))-1
WCSALE	Working capital to Sales	WCAP/SALE
CASALE	current asset to Sales	ACT/SALE
CSIS	cash and short-term investments/Sales	CHE/SALE
Liquidity		
WCTA	Working capital/Total Assets	WCAP/AT
CSIAT	Cash and short term investment to Total Assets	CHE/AT
Solvency		
CACL	current assets to current liabilities	ACT/LCT
QCACL	(current assets - inventory) to current liabilities	(ACT-INVT)/LCT

Table V: Univariate analysis

Ratio	A Priori	SMEs		Micro		Small		Medium	
		β	**Chi^2**	**β**	**Chi^2**	**β**	**Chi^2**	**β**	**Chi^2**
Profitability									
EBIDTAIE	(-)	-.0108494***	22.07***	-.063638***	13.09***	-.040803***	10.83***	-.022254***	21.75***
EBIDTATA	(-)	-1.19361***	19.50***	-.4889101***	4.72***	-.7048448***	4.75***	-2.107552***	27.01***
NISALE	(-)	-.2319702***	22.34***	-.1565969***	5.27***	-.0593781	0.39	-.2997896***	4.97***
RETA	(-)	-.1839748***	100.92***	-.0230958	0.90	-.0152018	0.12	-.3237676***	31.43***
NITA	(-)	-1.177601***	97.26***	-1.101897***	23.37***	-1.061895***	18.52***	-2.552399***	75.71***
NITE	(-)	.0511206	0.58	.0544052	0.34	.1791202	1.55	-.1472788	1.11
Leverage									
XINTTA	(+)	18.77527***	174.76***	6.154547***	10.07***	21.7166***	55.38***	36.05052***	121.76****
CLTA	(+)	2.587964***	196.91***	.8571542***	12.31***	2.504709***	40.24***	5.487994***	145.48***
TCTA	(+)	3.449214***	47.91***	1.690175***	6.46***	1.998587	3.56	5.395494***	21.88***
TLTA	(+)	2.280608***	296.29***	.7880578***	21.61***	2.826424***	97.01***	4.703341***	188.93***
STDEBV	(+)	.2985613	4.60	.098034	0.17	.3659389	1.73	.1431478	0.34
TDTA	(+)	.6552854	3.30	.3326899	0.46	1.955847***	8.03***	2.593609***	11.69***
Activity									
CETL	(-)	-.3459681***	107.01***	-.1571821***	18.95***	-.8830357***	49.42***	-1.592596***	88.91***
CASALE	(-)	-.3063103***	20.64***	-.376093***	18.36***	-.6182409***	14.16***	-1.187548***	20.33***
TTA	(-)	-15.99513***	143.34***	-26.546821***	25.76***	-27.16891***	52.25***	-24.35967***	130.08***
CG	(-)	-.4290882***	45.41***	-.1534083	3.03	-.2201675	3.13	-1.071719***	39.32***
WCSALE	(-)	-1.053956***	88.47***	-.3517744	8.51	-1.186603***	25.97***	-3.96608***	97.55***
CSIS	(-)	-.4340475***	23.30***	-.3558847	10.11	-1.460792***	25.74***	-1.658405***	20.76***
Liquidity									
WCTA	(-)	-2.199152***	197.37***	-1.2612188***	46.63***	-2.480588***	55.98***	-5.438686***	171.54***
CSIAT	(-)	-.7235075***	9.41***	-1.63	2.50	-5.79959***	46.76***	-5.173211***	38.79***
Solvency									
CACL	(-)	-.321967***	96.34***	-.1010271***	6.98***	-.4620135***	33.67***	-1.090778***	80.92***
QCACL	(-)	-.33133***	84.50***	-.1234249***	8.62***	-.1387676	3.55	-1.219197***	68.80***

This table reports the coefficients obtained from univariate regression analysis of respective covariates for different SMEs segments as discussed in section (3). For each size segment the coefficients estimated using discrete-time duration-dependant hazard function. *, **, * indicates that the coefficient is significant at the 1%, 5%, and 10% respectively.**

Table VI: Correlation matrix

Variable	EBIDTAIE	EBIDTATA	CACL	NITA	XINTTA	TLTA	CETL	TTA	CLTA	TCTA	CASALE	WCTA
EBIDTAIE	1											
EBIDTATA	0.5489*	1										
CACL	0.0412*	0.0916*	1									
NITA	0.4787*	0.9104*	0.1902*	1								
XINTTA	-0.1130*	-0.2401*	-0.4404*	-0.3283*	1							
TLTA	-0.1574*	-0.3471*	-0.6374*	-0.4305*	0.6936*	1						
CETL	0.0579*	0.0917*	0.8314*	0.1799*	-0.4925*	-0.7390*	1					
TTA	0.4465*	0.4049*	0.0960*	0.3461*	-0.1689*	-0.1676*	0.0630*	1				
CLTA	-0.1413*	-0.3832*	-0.6458*	-0.4480*	0.5213*	0.8266*	-0.6728*	-0.1324*	1			
TCTA	-0.1045*	-0.3058*	-0.4980*	-0.3429*	0.3504*	0.4871*	-0.5189*	-0.0831*	0.7051*	1		
CASALE	-0.3391*	-0.4079*	0.4832*	-0.2947*	-0.1543*	-0.2187*	0.4098*	-0.1923*	-0.2146*	-0.2422*	1	
WCTA	0.1315*	0.2828*	0.7789*	0.3745*	-0.5500*	-0.7577*	0.6040*	0.1956*	-0.7382*	-0.4937*	0.3055*	1

This table lists the correlation matrix among the covariates used in this study. The * indicates that the correlation is significant at the 1%.

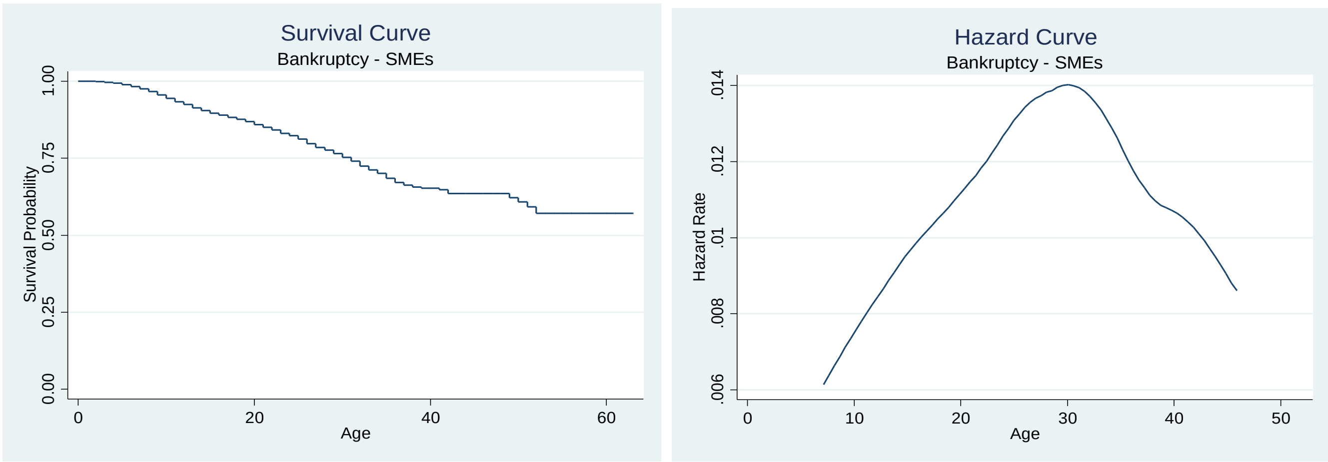
Table VII: Descriptive statistics

Variable		Micro		Small		Medium		SMEs	
		Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
EBIDTAIE									
	Failed	-11.393	21.059	-5.073	13.867	1.601	13.975	-4.518	18.287
	Healthy	-7.216	26.241	-4.178	31.029	10.183	34.941	1.073	33.227
NITA									
	Failed	-0.585	0.498	-0.390	0.449	-0.231	0.345	-0.347	0.457
	Healthy	-0.301	0.519	-0.269	0.406	-0.073	0.255	-0.228	0.412
TLTA									
	Failed	0.760	0.520	0.948	0.413	0.911	0.396	0.840	0.476
	Healthy	0.686	0.522	0.535	0.382	0.484	0.293	0.545	0.390
TCTA									
	Failed	0.153	0.135	0.132	0.115	0.122	0.108	0.130	0.124
	Healthy	0.141	0.129	0.115	0.099	0.101	0.084	0.114	0.101
TTA									
	Failed	0.005	0.021	0.001	0.017	0.006	0.023	0.004	0.021
	Healthy	0.014	0.017	0.008	0.024	0.018	0.030	0.012	0.027
CASALE									
	Failed	0.986	1.021	0.767	0.852	0.496	0.506	0.806	0.896
	Healthy	1.425	1.227	1.150	1.106	0.812	0.823	1.045	1.039
WCTA									
	Failed	-0.006	0.442	-0.007	0.347	0.046	0.340	0.021	0.399
	Healthy	0.047	0.440	0.268	0.358	0.322	0.280	0.245	0.362

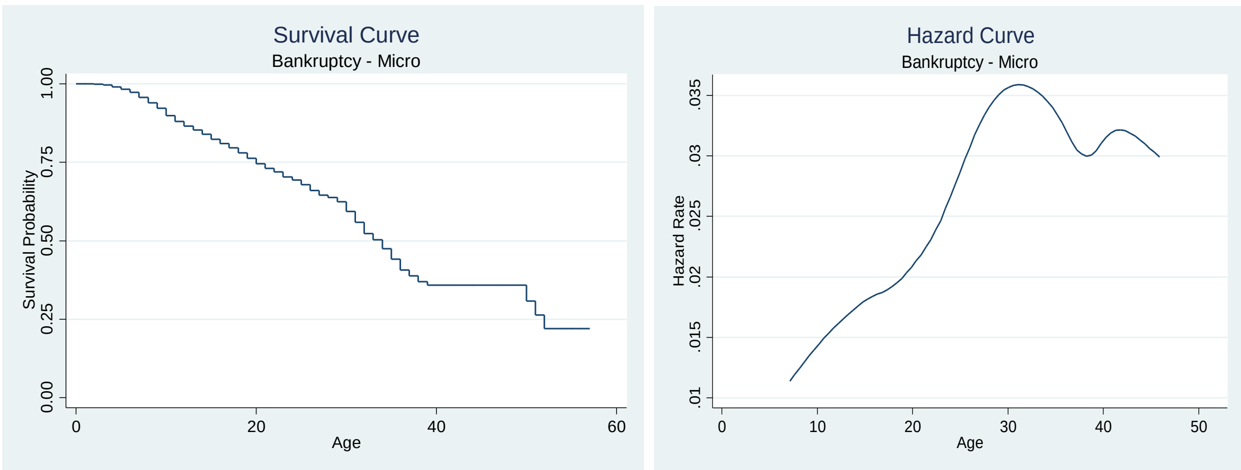
This table reports the descriptive statistics of the independent variables used in the study followed by the failed and healthy groups in the second column. The statistics are provided for the whole SME sample, micro, small, and medium.

Figure I: Survival and Hazard Curves

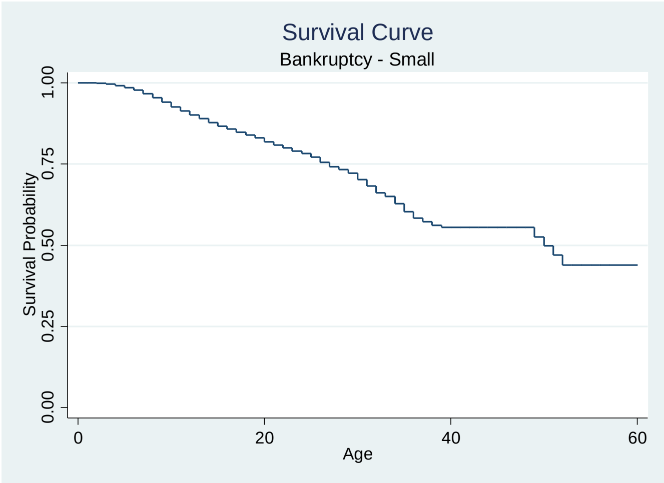
A. SMEs Survival and Hazard curves

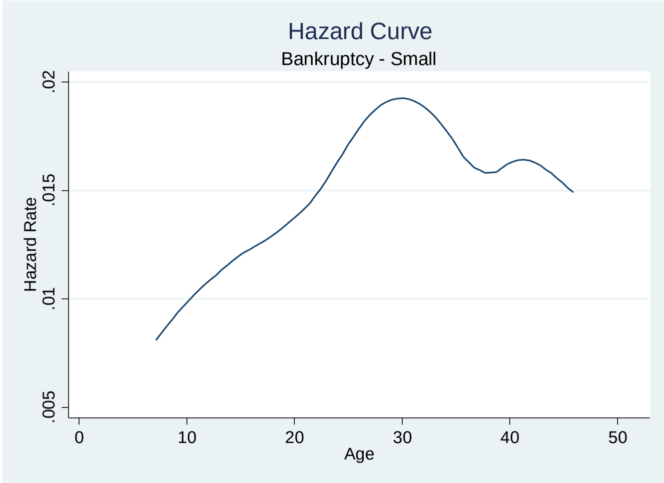


B. Micro SMEs survival and Hazard curves

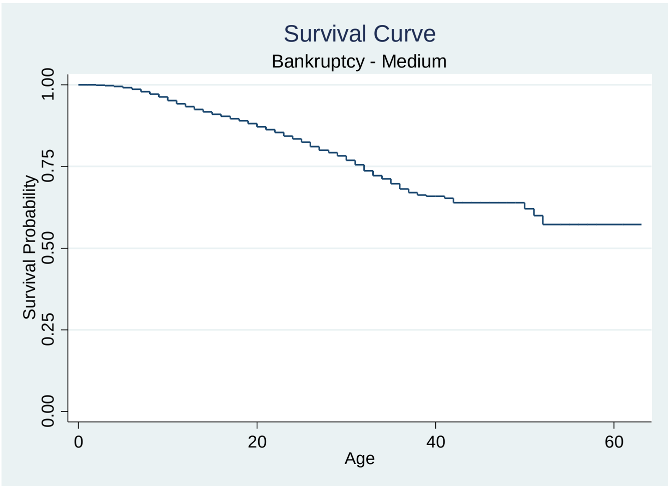


C. Small SMEs survival and Hazard curves





D. Medium SMEs survival and Hazard curves



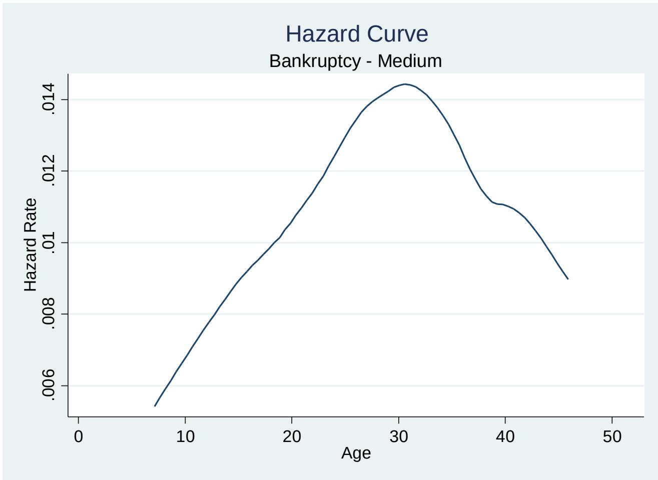


Table VIII: Multivariate Hazard Models

		Micro		**Small**		**Medium**		**SMEs**	
Variable		Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
EBIDTAIE	28	-0.0025	0.0048	-0.0021	0.0072	-0.0130*	0.0073	-0.0109***	0.0032
NITA	5	-0.0924	0.2172	-0.0943	0.3574	-0.3329	0.4290	-0.1698	0.1576
TLTA	31	1.34949***	0.3033	2.8226***	0.4429	2.6630***	0.4647	1.9519***	0.2043
TCTA	33	0.9417	0.9348	3.1230***	1.3415	2.8267**	1.4312	1.2203**	0.6225
WCTA	20	-1.186459***	0.0010	-0.6447	0.5101	-3.3204***	0.5819	-0.6221***	0.2444
TTA	32	-2.0863	4.5838	-15.5357***	7.1535	-11.9163***	5.4457	-7.1121***	2.7305
CASALE	24	-0.3677	0.1047	-0.2871	0.1842	-0.3697	0.2538	-0.1110	0.0761
Constant		-12.7023***	1.7318	-14.6922***	2.2521	-13.0451***	1.7266	-13.1424***	1.3608
IRC		0.2851***	0.0544	0.3595***	0.0790	0.2243***	0.0671	0.1832***	0.0327
Age dummies		Yes		Yes		Yes		Yes	
Industry control		Yes		Yes		Yes		Yes	
Goodness of fit									
Wald chi2		157.0300***		167.63***		220.9000***		549.9100***	
Log Likelihood		-772.9203		-806.1630		-1124.5943		-2933.9049	
AIC		5991.81		2367.189		1720.326		1641.84	
BIC		6567.009		2822.551		2179.792		2029.234	
Number of observations		16,614		23,640		36,630		79,016	

This table reports the estimations corresponding to micro, small, medium, and SMEs respectively. For each segment the table reports the results obtained from respective multivariate hazard analysis followed by goodness of fit measures. *, **, * indicates the significance at the 1%, 5%, and 10% respectively.**

Table IX Classification performance measure

Decile	Micro	Small	Medium	SMEs
1	18.67%	23.67%	21.45%	24.41%
2	25.51%	24.00%	27.65%	31.00%
3	19.00%	17.33%	15.75%	13.73%
4	9.51%	14.55%	12.45%	8.33%
5	8.66%	8.32%	9.05%	5.00%
6 , 10	18.65%	12.13%	13.65%	17.53%
Total	100.00%	100.00%	100.00%	100.00%

This table reports the classification performance measures for each of the SMEs size segments: Micro, small, medium, and whole SMEs sample for the period from 2009 till 2013. Values in parenthesis are cumulative classification measures over the ten deciles.