PUBLIC VS PRIVATE SMES: A COMPARISON OF DISTRESS HAZARD

Jairaj Gupta
Department of Finance and Economics, Brighton Business School,
Brighton, BN2 4AT, UK; email: jairajgupta@outlook.com

Andros Gregoriou
Department of Finance and Economics, Brighton Business School,
Brighton, BN2 4AT, UK; email: A.Gregoriou@brighton.ac.uk

Gbenga Ibikunle
Department of Accounting and Finance, University of Edinburgh Business School,
Edinburgh, EH8 9JS, UK; email: Gbenga.Ibikunle@ed.ac.uk

April, 2015

Abstract

This study considers listed and unlisted small and medium-sized enterprises (SMEs) of the United States separately while developing one-year financial distress prediction model for them. Empirical analysis of financial distress performed using discrete-time duration-dependent hazard rate modelling technique with logit link and a set of financial covariates reveal striking differences between distress hazard of listed and unlisted SMEs. Almost an identical set of covariates exhibit significant discriminatory power for both listed and unlisted SMEs, but there exist significant differences in their weights of regression coefficients in respective groups. Further, Average Marginal Effects of respective covariates for unlisted group of SMEs are strikingly higher than their listed counterparts, suggesting higher vulnerability of unlisted firms due to changes in financial ratios. Our findings support the view that stock exchange listing can relieve SMEs from external financing constraints, thus reducing their likelihood of financial distress.

Keywords: Financial Distress; small and medium-sized enterprises; Discrete Hazard Models; Liquidity; Credit Risk

JEL Classification Codes: G12; G32; G33

1. Introduction

Small and medium-sized enterprises (SMEs) are widely considered to be a fundamental component of an economy, and are viewed as an important route to recovery in the aftermath of the global financial crisis of 2008-2009. Given the increased importance of SMEs, a significant volume of academic literature on SMEs financial distress has emerged in recent years (e.g. Altman and Sabato 2007, Gupta, Wilson, *et al*. 2014a, 2014b, Keasey *et al*. 2014). Among several reasons, access to external finance is unanimously the most important factor hindering SMEs growth, development (e.g. Beck and Demirguc-Kunt 2006, Ardic *et al*. 2012) and potentially, survival. Lack of collateral and information asymmetries reduces their access to bank financing, while stock exchange listing could relieve them from financing constraints (Kim 1999). Thereby, they may relax their overdependence on lending institutions/banks for external financing by listing themselves in stock exchanges, consequently removing the financial barriers hindering their growth and competitiveness. However, listing might be difficult due to admission requirements and disclosure regulations (see Gao *et al*. 2013). This realization has led to the emergence of stock markets with relaxed admission requirements and disclosure regulations specifically targeting SMEs (e.g. Alternative Investment Market of the London Stock Exchange). Disclosures can reduce information asymmetry between firms and external financers/investors, which in turn can make access to external finance easier. As a consequence, listed SMEs are expected to experience lower financial distress hazard than their unlisted counterparts.

We contribute to the literature on SMEs by examining if there are significant differences in the determination of financial distress of listed and unlisted SMEs. Our empirical question is motivated by the Information Cost Liquidity Hypothesis (ICLH) in the market microstructure literature. The ICLH was first established by Van Horne (1970) in the context of new listings on the New York Stock Exchange, stating that listing signals good news about firms' future prospects. Since the work by Horne (1970), researchers such as Shleifer (1986), Dhillon and Johnson (1991), Beneish and Gardner (1995), Hegde and McDermott (2003), Gregoriou and Ioannidis (2006), Liu (2011) and Gregoriou (2011) have examined whether information about the investment appeal of a stock is provided by news of listing changes. They all report significant improvement in stock's performance after inclusion in the index. Therefore, considering the previous literature we expect listed SMEs to be more profitable and less susceptible to financial distress than their unlisted counterparts. Furthermore, if as a consequence of listing SMEs possess a richer information environment; trading may be more frequent, resulting in increased liquidity. In order to empirically test this hypothesis we include proxies for liquidity as factors in explaining financial distress of listed SMEs. In particular, we use liquidity ratio of Amihud (2002) and illiquidity metric of Florackis *et al*. (2011) as suitable proxies.

We empirically test our hypothesis using sample of listed and unlisted SMEs of the United States covering sampling period between 1980 and 2014. Firm level annual accounting information is sourced from Compustat and monthly stock prices from CRSP databases. Considering the suggestion of Gupta *et al*. (2015), we use *discrete-time duration-dependent hazard model with logit link* to perform univariate and multivariate one-year financial distress hazard analysis of listed and unlisted SMEs respectively. Financial ratios with established reputation of distress prediction in earlier studies are being used as covariates along with liquidity measure of Amihud (2002) and illiquidity measure of Florackis *et al*. (2011). Our definition of financial distress based on firms' financial performance is adapted from Keasey *et al*. (2014). To gauge within-sample classification and out-of-sample validation performance of multivariate models developed, we estimate area under the *Receiver Operating Characteristic* (ROC) curves of respective hazard models[1].

Based on our empirical findings, we report significant differences between distress hazard of listed and unlisted SMEs. In univariate analysis, an identical set of financial ratios are significant in discriminating between financially distressed and censored group of listed and unlisted SMEs, but we observe statistically significant difference in weights of regression coefficients of respective covariates (except tax/total assets) of listed and unlisted SMEs. *Average Marginal Effects* (AME) of respective covariates for unlisted group of firms are strikingly higher than their listed counterparts, suggesting higher vulnerability of unlisted firms due to changes in financial position. For listed SMEs, both liquidity factors also exhibit significant discriminatory power. Findings from our multivariate analysis also resonate with our univariate findings. STDEBV (short term debt/equity book value) and TCTA (trade creditors/total assets) are significant predictors of distress hazard of listed SMEs, but are excluded from the multivariate hazard model developed for unlisted SMEs. Suggesting listing can improve firms' access to external finance. Further, estimated regression coefficients of mutual covariates in respective distress prediction models for listed and unlisted SMEs also show striking differences in their weights. This is validated by performing Wald test of

[1] See Hosmer Jr *et al*. (2013) for relevant discussion on suitability of ROC curves in measuring classification performance of binary prediction models.

equality of regression coefficients of mutual covariate. Out of five mutual covariates, three exhibit significant statistical difference in the weight of the regression coefficients of listed and unlisted SMEs. This is further reinforced when we compare AME of respective covariates for hazard models of listed and unlisted groups. Similar to univariate analysis, AME of mutual covariates are significantly higher for unlisted SMEs than their listed counterparts. This supports our hypothesis that listed SMEs are less vulnerable to changes in financial position than unlisted SMEs. Additionally, the illiquidity measure based on Amihud (2002) enter significantly in the multivariate setup but the measure based on Florackis *et al*. (2011) is weakly significant. Finally, the within-sample and hold-out-sample area under ROC curves for all our multivariate distress hazard models are above 0.80, indicating very strong classification performance of our distress prediction models. Our results can be attributed to the ICLH given that the greater the liquidity of SMEs, the lower the probability of financial distress. This suggests that investors could demand a lower premium for holding stocks with relatively more available information.

The remainder of the paper is structured in the following way: the next section defines our liquidity measures; section 3 outlines our empirical methods including an explanation of the dataset. Empirical results are reported in Section 4 and finally section 5 concludes our findings.

2. LIQUIDITY MEASURES

There are various measures of liquidity used in the academic literature. A vast majority of research papers use the bid-ask spread to capture liquidity effects around the announcement of news. In order to capture the financial stability of a firm, the Amihud (2002) ratio is used in previous market microstructure studies. We therefore compute the Amihud (2000) ratio by estimating Equation (1) in the following way.

$$RtoV_{it} = \frac{1}{M_{it}} \sum_{m=1}^{M_{it}} \frac{|R_{itm}|}{V_{itm}} \tag{1}$$

where, R_{itm} and V_{itm} are, respectively, the return and the monetary volume of stock i on month m at year t, and M_{it} is the number of valid observation months in year t for stock i.

As argued by Florackis *et al*. (2011), the $RtoV$ ratio cannot compare stocks with different market capitalization and therefore carries a significant size bias, i.e. small cap stocks are bound to exhibit lower trading volume (in monetary terms) than big cap stocks leading to a size bias. Under the Amihud (2002) $RtoV$ ratio (Equation 1), small cap stocks are automatically characterized as ‘‘illiquid’’ due to their size. Therefore, for robustness purposes we also calculate the Florackis *et al*. (2011) $RtoTR$ ratio by computing the following equation:

$$RtoTR_{it} = \frac{1}{M_{it}} \sum_{m=1}^{M_{it}} \frac{|R_{itm}|}{TR_{itm}} \tag{2}$$

where, $RtoTR_{it}$is the turnover ratio of stock i on month m, and M_{it} and R_{itm} are defined as previously. The *RtoTR* ratio is free from size bias as there is no empirical association between turnover and market value (for more information see Florackis *et al*. (2011)). It should be noted that both our liquidity measures are actually encapsulating illiquidity because the trading intensity variable is expressed in the denominator of each ratio.

3. EMPIRICAL METHODS

This section provides discussion related to the source and use of dataset, selection of explanatory variables and statistical models that we use for our analysis.

3.1 DATASET

We sourced firm-level annual accounting information of the United States SMEs from Compustat and monthly stock prices of listed SMEs from CRSP databases. We consider a firm as SME if it reports annual sales turnover of less than $65 million. Considering the significant changes that were introduced in the Bankruptcy Reform Act of 1978, we ignore bankruptcy filings prior to 1980 and choose firms that filed for bankruptcy between January 1980 and December 2014[2]. However, in this study we concentrate on financial distress rather than legal bankruptcy with the presumption that it is the primary reason behind bankruptcy and always precedes the bankruptcy filing event. Further, filing for legal bankruptcy is the least efficient exit strategy for SMEs (Balcaen *et al*. 2012) and distress definitions based on bankruptcy laws are inefficient in comparison to distress definition based firms' financial performance (see Gupta *et al*. 2015). Thus, following Keasey et al. (2014), an SME experiencing financial distress is defined as one that satisfies the following: (i) its expenses exceeds earnings during two consecutive years, (ii) its total debt exceeds net worth during

[2] However, after applying all required filters and excluding observations with missing values our sampling period narrows down between 1982 and 2014.

those two years in (i), and (iii) it records negative growth in net worth during the same consecutive periods in (i) and (ii). Additionally, a firm is also recorded as financially distressed in the year immediately following these distress events.

We proxy a firm's age as the earliest year, for which annual financial information is available for that firm in the Compustat database. 1950 is the earliest data entry year for firms' financial information in Compustat, thus the maximum age that a firm could have is 64 years. Furthermore, firms with Standard Industrial Classification (SIC) codes from 6,000 through 6,999 (financial firms) and 4900 through 4949 (regulated utilities) are excluded from the sample. We also exclude subsidiary firms (if 'stock ownership code' (Compustat data item 'stko') is '1' (subsidiary of a publicly traded company) or '2' (subsidiary of a company that is not publicly traded) in the Compustat database). We consider a SME as listed if it is publicly traded in any of the three popular exchanges, i.e. NYSE, AMEX and NASDAQ (Compustat data item 'exchg' is 11 (NYSE), 12 (AMEX) or 14 (NASDAQ)) and unlisted otherwise (Compustat data item 'exchg' is 1 (non-traded company), 13 (OTC Bulletin Board) or 19 (Other OTC)). Consequently, the final dataset consists of 40,078 firm-year observations with 11,719 records for listed and 38,359 records for unlisted US-based SMEs[3] (see Table 1).

[Insert Table 1 Here]

3.2 Selection of Covariates

To develop the hazard models we employ financial ratios that are already established as significant predictors of SMEs default risk. The adopted covariates assess firms' performance on liquidity, solvency, activity, profitability and interest coverage dimensions. Specifically, we incorporate the covariates from popular studies on SMEs bankruptcy such as Altman and Sabato (2007), Lin *et al*. (2012), Gupta *et al*.(2014) and others[4]. Two of the adopted covariates in this paper are novel to the SMEs default risk literature. These are measures of long-term illiquidity or liquidity. The first is the Amihud (2002) ratio, defined as the ratio of the absolute return to trading volume. However, trading volume is expected to be materially larger for high transaction instruments, thus leading to a large firm bias. As a result, we also compute the Florackis *et al*. (2011) illiquidity metric as a robustness test. This is because in

[3] Note that given these exclusion criteria, firms could have multiple entry and exits in the dataset. For example, when an existing SME reports annual sales revenue over $65 million it exits our sample and returns only when its sales revenue drops below $65 million.
[4] Altman *et al*. (2010) and Gupta *et al*. (2014) provide detailed discussions of the covariates selected as well as their relationship with the probability of a default.

the Florackis *et al*. (2011), trading volume is replaced by turnover which is free from size bias. All the covariates along with their respective definitions are listed in Table 2.

[Insert Table 2 Here]

3.3 Hazard Model

3.3.1 Basic Hazard Model

The survival analysis conducted in this study involves estimation of the time duration taken for an event to occur; in this case the event is a firm experiencing financial distress. Suppose T is a non-negative random variable denoting the time to a distress event and *t* corresponds to the survival of a firm beyond time *t*. Choosing not to express *T*s probability density function as $f(t)$ or its cumulative distribution function (CDF) as $F(t) = \Pr(T \leq t)$, rather envisaging *T*s survivor function, $S(t)$ or its hazard function $h(t)$ significantly simplifies the survival analysis concept (Cleves *et al*. 2010). The survivor function estimates the probability of survival beyond time *t*, which is essentially the inverse CDF of *T*, i.e.:

$$S(t) = 1 - F(t) = \Pr(T > t) \tag{3}$$

At $t = 0$ the survivor function equals one and approaches zero as t advances towards infinity. The relationship between the survivor function and the hazard function (also referred as the conditional failure rate at the time t) can be expressed mathematically defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} = \frac{-d \ln S(t)}{dt}; \tag{4}$$

Unambiguously, the hazard rate is defined as the (limiting) probability that the failure event occurs within a set time interval, given that the subject has survived to the onset of that time interval, divided by width of the time interval. The hazard rate takes on values from zero to infinity and may increase, decrease or remain constant over time. A hazard rate of zero indicates no risk of failure during the period in which it is computed, while infinity signifies certainty of failure.

3.3.2 Discrete Hazard Model

An event occurring at any instant in continuous-time such that the exact censoring and survival times are recorded in relatively fine time scales like seconds, hours or days and there are no *tied* survival time periods, then one may employ continuous-time survival model in computing the likelihood of an event’s occurrence (Rabe-Hesketh and Skrondal 2012).

However, if the data has relatively few censoring or survival times with *tied* survival time periods, the discrete-time survival model is considered a more appropriate choice where coarse times-scales are commonly used, for example, expressing time to event in weeks, months or years (Rabe-Hesketh and Skrondal 2012). Interval-censoring[5] results in discrete-time data, which is the case with our sample. Here, the start and end of each time interval is same for all the SMEs in the analysis time, given that the information provided is recorded on an annual basis. Therefore, the event of interest may take place at any time within the year; however, the detail is not available until the year concludes.

Based on the foregoing, we therefore estimate our hazard models in discrete-time framework with *random effects* (α_i), thus controlling for *unobserved heterogeneity* or *shared frailty*. The discrete-time representation of the continuous-time proportional hazard model with time-varying covariates results in a generalized linear model with *complementary log-log* (Grilli 2005; Jenkins 2005; Rabe-Hesketh and Skrondal 2012) link, expressed in the following way:

$$cloglog\big(h_i(t)\big) \equiv \ln\{-\ln(1 - h_i(t))\} = \beta x(t)_i' + \lambda_t \tag{5}$$

λ_t denotes the time-specific constant estimated freely for each time period *t*, thus no assumption is made about the baseline hazard function within the specified time interval. However, in most empirical studies logit link is used over complementary log-log (clog-log) link as specified in Equation 6.

$$P_{i,t} = \frac{e^{\alpha(t)+\, x(t)_i^{'}\beta}}{1 + e^{\alpha(t)+\, x(t)_i^{'}\beta}} \tag{6}$$

Where *α(t)* captures the baseline hazard rate and $P_{i,t}$ corresponds to the probability that subject *i* experiences the event at time *t*. This should yield strikingly similar results if the time intervals are small (Rabe-Hesketh and Skrondal 2012) and the sample bad rate (% of failed to non-failed) is very low (Jenkins 2005). One may also select a probit link function, assuming there are compelling reasons to that the underlying distribution of the process being modelled is normal, or if the event being studied is not a binary outcome but a proportion (e.g. proportion of population at different income levels). While these specifications will generally produce results that are quite similar, there are significant differences in terms of non-proportionality (see Sueyoshi (1995) for detiled discussion). Therefore, considering this

[5] The event in this case occurs in continuous-time; however, we only record the time interval within which the event takes place.

discussion and the suggestion by Gupta *et al*. (2015), we employ a discrete hazard model with logit link to develop our distress prediction models.

3.3.3 Specification of Baseline Hazard Rate

The next step is the specification of baseline hazard function, the hazard rate when all the covariates are set to zero. This stage in the analysis precedes the estimation of the discrete-time hazard model itself. The specification of the baseline hazard function is achieved by defining time-varying covariates that exhibit functional relationships with survival times. Established specifications include log(survival time), polynomial in survival time, fully non-parametric and piece-wise constant (Jenkins 2005). Fully non-parametric baseline hazard specification requires generation of age specific dummy variables. Assuming no missing time spells, the number of dummies should correspond to the maximum survival time in the dataset. For example, if the upper limit of survival time equals fifty years, fifty dummies are needed for model estimation[6] (see for example, Beck *et al*. 1998). This approach is complicated by lengthy survival times as is the case with firm bankruptcy. A convenient alternative for specifying the baseline hazard function is to employ piece-wise constant method. This approach involves splitting the survival times into different time intervals that are assumed to exhibit a constant hazard rate (Jenkins 2005). However, duration specific hazard rate cannot be estimated for intervals/dummies with no events (see Jenkins 2005; Rabe-Hesketh and Skrondal 2012). Given its estimation convenience, the piece-wise constant specification of baseline hazard rate is desirable. Notwithstanding, it should be noted that the if the hazard curve shows recurrent and continuous sharp rises and falls, the fully non-parametric baseline hazard shall be a more appropriate choice.

3.4 Performance Evaluation

In order to assess the performance of the models developed, we report area under the ROC curves, which is a widely used non-parametric method of evaluating model's classification performance (see Hosmer Jr *et al*. 2013). The ROC curve is obtained by plotting the true positive (when a firm actually defaults and the model classifies it default status as expected) against the false-positive (when a firm does not default but the model classifies its default status as expected) rate as the threshold to discriminate between non-failed and failed firms' changes. The area under ROC curves (AUROC) encapsulates the precision of the model; with AUROC equalling 1 denoting a model with perfect prediction accuracy and equalling 0.5

---

[6] The model is estimated using forty nine dummies in order to avoid perfect multicollinearity arising from the dummy variable trap.

suggest no discrimination (see Anderson, 2007). Although there is no 'golden rule' regarding the value of AUROC, however anything around 0.8 or above is considered to be excellent. The Gini coefficient and Kolmogorov–Smirnov (K–S) statistics, usually employed as gauges of the performance of a scoring model, can be easily computed from AUROC. The Gini coefficient, defined as G = 2(AUROC − 0.5), captures the consistency in the prediction of the model as developed, while the K–S statistics quantifies the distance between the failed and non-failed distributions at the optimal cut-off point and is about 0.8 × Gini coefficient. A model with K–S statistics value below 20 should be re-examined, likewise a model having a value above 70 is probably too good to be true and should also be re-examined (see Anderson 2007).

4. RESULTS AND DISCUSSION

We commence our discussion of the results by considering the descriptive statistics of the covariates along with the extent of correlation among them. This follows by univariate hazard analysis of each covariate in turn by using financial distress definition as discussed earlier in section 3.1. Then we discuss the development of multivariate discrete-time duration-dependent hazard models based on Average Marginal Effects (AME) of respective covariates along with the baseline hazard specification. We also illustrate the steps involved in developing various multivariate hazard models along with relevant analysis relating within-sample and out-of-sample classification performance of respective multivariate hazard models. To eliminate the influence of extreme outliers on our statistical estimates, the range of all financial ratios employed is restricted between 5th and 95th percentiles. Following the suggestion of Gupta, Gregoriou, *et al*. (2014), we also employ dummy indicators for micro (annual sales revenue is less than \$ 2.6 million) and small (annual sales revenue is greater than \$ 2.6 million but less than \$ 13 million) firms into our hazard analysis to account for any differences that may arise due to firms' size. Finally, all covariates are lagged by one-time period in order to ensure that the information is available at the beginning of each time period.

4.1 DESCRIPTIVE STATISTICS AND CORRELATION

Initial inspection of descriptive statistics is useful in evaluating the variability of the covariates and the potential biasness that may arise in the multivariate setup due to unexpected extreme fluctuations. We expect the mean of covariates that exhibit positive relationship with the insolvency hazard to be higher for the distressed group than for healthy or censored group (e.g. see the variable TLTA in Table 3) of firms. On the contrary, the mean of covariates that shows negative relationship with the insolvency hazard is expected to be lower for the default group than for their healthy counterparts (e.g. see variable CETL in Table 3). Mean, median and standard deviation of all covariates are as per our expectation for respective group of listed and unlisted SMEs, except CTA (for listed SMEs) and STDEBV (for unlisted SMEs). These covariates might be problematic in the multivariate setup.

[Insert Table 3 Here]

The correlation matrix presented in Table 4 provides evidence that some of the covariates are strongly correlated with each other. For example, FETA exhibits moderate to strong correlation with six other covariates. This is also the case with TCTA and LCR, while RETA shows strong positive correlation of approximately 0.74 with EBITDATA, supporting the expectation that SMEs primarily rely on internal sources for their funding requirements. In order to address this issue of multicollinearity effectively while developing multivariate models, we use a selection procedure of covariates based on their Average Marginal Effects obtained from the univariate analysis. Detailed discussion on this will follow soon. Moreover, casual observation of the means of respective covariates for listed and unlisted group of SMEs reveal striking differences expect TTA. Thus we get initial motivation to believe that the weights of the regression coefficients might be different for listed and unlisted firms.

[Insert Table 4 Here]

4.2 UNIVARIATE ANALYSIS OF COVARIATES

We estimate univariate discrete hazard models for respective covariates in turn using Equation 6 separately for listed and unlisted firms. In both cases the dependent variable has binary outcome, where ‘1’ implies the firm has experienced the financial distress event and ‘0’ otherwise or censored. As evident from the estimated results reported in Table 5, all covariates are highly significant in discriminating distressed and censored firms for both groups with expected sign of respective coefficients, except STDEBV for unlisted SMEs. We expect the sign of it’s coefficient to be positive but it’s negative, which might be due to the lower mean of STDEBV for distressed group than censored ones. Furthermore, we find evidence that both our liquidity measures are positive and significant for listed SMEs. This implies that the probability of default is related to a lack of liquidity of the stocks. This provides evidence of the ICLH given that investors prefer not to hold securities with less information causing the probability of financial distress to increase.

We also see in Table 5 that weights of regression coefficients of respective covariates for listed and unlisted groups are strikingly different. Casual comparison of their *Average Marginal Effects*[7] (AME; dy/dx) reinforces our hypothesis, as we see large differences in AME of respective covariates between listed and unlisted groups. For respective covariates, the AME is significantly higher for unlisted SMEs that their listed counterparts. This suggests that default probabilities of listed SMEs are less affected by unit change in value of respective covariate than unlisted SMEs. Overall, unlisted SMEs seem to be more vulnerable to financial distress due to changes in their financial ratios than listed SMEs. This supports our hypothesis that listed SMEs are less susceptible to financial distress than unlisted ones. In order to statistically test the differences in the weights of regression coefficients of respective covariates in different groups, we use '-gsem-' command in Stata 13. It performs the Wald test of equality of coefficients of mutual covariates obtained from two different regression models. The p-values of this test are reported in the last column of Table 5, which shows highly significant statistical difference in the weights of the regression coefficients of all respective covariates except TTA. This strongly suggests that although the default attributes for both listed and unlisted SMEs are mutual but they need to be treated separately while developing credit risk models. Further, in line with our hypothesis both liquidity measures *RtoV* and *RtoT* are highly significant in discriminating between distressed and censored listed SMEs.

[Insert Table 5 Here]

4.3 Developing Multivariate Hazard Models

We start this section with baseline hazard specification based on Kaplan-Meier estimates of hazard curves (see Cleves *et al*. 2010), followed by development of multivariate discrete-time duration-dependent hazard models with logit link for our sample of listed and unlisted SMEs. The dependent variable for both these models has binary outcome with financially distressed

[7] In non-linear regression analysis, Marginal Effects is an useful way to examine the effect of changes in a given covariate on changes in the outcome variable, holding other covariates constant. These can be computed as marginal change (it is the partial derivative for continuous predictors) when a covariate changes by an infinitely small quantity and discrete change (for factor variables) when a covariate changes by a fixed quantity. Whereas, Average Marginal Effects (AME) of a given covariate is the average of its marginal effects computed for each observation at its observed values. Alternatively, AME can be interpreted as the change in the outcome (financial distress = 1; in our case) probabilities due to unit change in the given covariate, provided other covariates are held constant. See Long and Freese (2014) for detailed discussion on this topic.

equalling ‘1’ and ‘0’ otherwise, while independent variables are the set of covariates found significant in the univariate regression analysis. Considering the multicollinearity among the covariates, we introduce each covariate in turn into the multivariate setup based on the magnitude (sign is ignored) of their AME. For this, at first we rank[8] all the covariates found significant in the univariate analysis based on the absolute value of their AME (see columns six and ten in Table 5) and then start introducing each covariate in turn into the multivariate setup in increasing order of the rank of their AME. The rational being, higher the value of AME, higher will be the change in the predicted probability due to unit change in the covariate. Thus a covariate with higher value of AME (e.g. FETA in Table 5) is more efficient in discriminating between distressed and censored firms than covariates with lower value of AME (e.g. TLTA in Table 5). Further, if the introduction of a covariate affects the sign[9] of any previously added covariate, then that covariate is excluded from the multivariate model. This can possibly happen due to multicollinearity among covariates, thus their exclusion seems to be a reasonable choice. Moreover, we believe that this method of covariate introduction while developing the multivariate models leaves us with best set of covariates with expected sign of coefficients of respective covariates. Additionally, we also control for volatile macroeconomic environment and varying distress rates across different time periods by introducing year dummies in the multivariate hazard models.

Final set of multivariate hazard models reported for both listed and unlisted SMEs are estimated using observation from entire sampling period available to us, thus we do not have separate test and holdout samples. In order to assess the within-sample classification performance of the models developed we estimate area under ROC curve for respective models using the full estimation sample (i.e. 1980 to 2014). For out-of-sample validation we first estimate multivariate hazard model using observation till the year 2011 and using these estimates we predict the default probabilities for the year 2012; then we include 2012 in the estimation sample and predict default probabilities for 2013 and so on, till the year 2014. Then we use the predicted default probabilities from the year 2012 through 2014 to estimate out-of-sample AUROC for respective multivariate hazard models.

[8] Highest value gets rank ‘1’, second highest gets rank ‘2’ and so on.
[9] Coefficients with negative sign become positive and vice versa.

4.3.1 DETECTION OF BASELINE HAZARD RATE

Figure 1 shows hazard curves for listed and unlisted SMEs estimated using Kaplan-Meier[10] estimator. As we see in Figure 1, the distress hazard of both listed and unlisted SMEs rises as firms get older, however for a given age this rise is almost double for unlisted SMEs than their listed counterparts. At the age of twenty the distress hazard of unlisted firms is almost 1, which implies certainty of failure. While for the same age the distress hazard is around 0.5 for listed SMEs. This implies that until the age of 20 years, unlisted SMEs are almost twice more vulnerable to financial distress than their listed counterparts. This difference in the hazard rates across the age category reinforce our hypothesis that listed and unlisted SMEs needs to be treated separately while modelling credit risk for them. Additionally, both hazard curves show steep rise with respect to firms' age. Under this situation fully non-parametric baseline hazard specification seems to an appropriate choice. Thus we include age specific dummies in our multivariate hazard models as specification for the baseline hazard rate.

[Insert Figure 1 Here]

[Insert Table 6 Here]

4.3.2 HAZARD MODEL FOR LISTED SMES

The multivariate hazard model estimated for listed SMEs is reported in Table 6. Considering our covariate introduction method as discussed earlier; out of thirteen significant covariates in the univariate analysis, we find nine covariates suitable for developing the multivariate hazard model for listed SMEs. All financial ratios other than TCTA are highly significant in discriminating between financially distressed and censored firms with significant AME. The within-sample AUROC is about 0.87 and out-of-sample AUROC is about 0.82, which emphasises excellent discriminatory performance of our multivariate hazard model in identifying distressed and censored firms (see Figure 2). The AME are reported in percentage, which states that TTA is the most powerful covariate with AME of around -20 followed by FETA with AME of around 10. The multivariate analysis is quantitatively similar to our univariate findings concerning the relationship between the probability of default and liquidity measures. *RtoV* and *RtoT* are significant respectively under 5% and 10% significance level. This is because like the univariate analysis, we find that stocks with less information are regarded as illiquid, resulting in a greater likelihood of bankruptcy for SMEs.

---

[10] See among others Cleves *et al*. (2010) and Mills (2011) for details on Kaplan-Meier hazard estimator.

4.3.3 HAZARD MODEL FOR UNLISTED SMES

The multivariate hazard model for unlisted SMEs is also reported in Table 6. As we see, out of eleven highly significant covariates in univariate analysis, seven are appropriate in the multivariate setup. We also see some differences in the factors affecting the default probability of listed and unlisted SMEs. For instance, STDEBV and TCTA are significant predictors of insolvency hazard of listed SMEs, but they do not find a place in the hazard model developed for unlisted SMEs. This suggests that listing lead to better access to external finance, which might turn out to be a significant reason for their finance distress. Unlike listed SMEs, CETL and LCR enter significantly in the multivariate hazard model for unlisted SMEs, which emphasises the importance of owners' equity on financial distress of unlisted SMEs. Further, Wald test of equality of regression coefficients of mutual covariate also show convincing results. Out of five mutual covariates, three exhibit significant statistical difference in the weight of the regression coefficients of listed and unlisted groups (see last column of Table 6). This is further reinforced when we compare AME of respective covariates for hazard models for listed and unlisted groups. As observed in the univariate analysis section, here also AME for all covariates are significantly higher for unlisted SMEs than their listed counterparts (see Table 6). This suggests that unlisted SMEs are more vulnerable to changing financial position unlike listed SMEs. Finally, the within-sample and hold-out-sample AUROC is about 0.85 (see Figure 2), which emphasises excellent classification performance of our multivariate hazard model developed for unlisted SMEs.

[Insert Figure 2 Here]

5. CONCLUSION

Access to external finance is unanimously the principal factor obstructing SMEs growth and development. This might be due to lack of collateral and information asymmetries. Prolonged difficulty in accessing external finance may lead to financial distress or bankruptcy. However, stock exchange listing could relieve SMEs from external financing constraints (Kim 1999). Consequently reducing their overdependence on banks for external financing and thereby, reducing their likelihood of financial distress.

We empirically test this hypothesis using sample of listed and unlisted SMEs of the United States covering sampling period between 1980 and 2014. One-year financial distress hazard analysis of listed and unlisted SMEs is performed using discrete-time duration-dependent hazard rate modelling technique and set of financial covariates with established significance of financial distress prediction in earlier studies. The definition of financial distress employed based on firms' financial performance is adapted from Keasey *et al*. (2014). We report significant differences between distress hazard of listed and unlisted SMEs. Although identical set of financial ratios are significant in discriminating between financially distressed and censored group of listed and unlisted SMEs, but we report statistically significant difference in the weights of regression coefficients of respective covariates (except TTA) of listed and unlisted SMEs. AME of respective covariates for unlisted group of firms are strikingly higher than their listed counterparts, suggesting higher vulnerability of unlisted firms due to changes in financial ratios. Additionally, regression coefficients of mutual covariates in multivariate hazard models for listed and unlisted SMEs also show striking differences in their weights. Three out of five mutual covariates exhibit significant statistical difference in the weight of their regression coefficients. Our hypothesis is further reinforced when we compare AME of respective covariates for hazard models of listed and unlisted groups of firms. In line with univariate analysis, AME of mutual covariates are significantly higher for unlisted SMEs than listed ones.

We also find that liquidity is a critical factor in explaining financial distress of listed SMEs, since we report positive association between default risk and lack of liquidity. Our results can be explained by the LCIH, where investors sell stocks with a poorer information environment contributing to an increased probability of bankruptcy of listed SMEs. Given the importance of SMEs and how the absence of liquidity contributed to the recent global financial crises, the results in our paper cannot be ignored.

References

Altman, E.I. and Sabato, G., 2007. Modelling credit risk for SMEs: Evidence from the US market. *Abacus*, 43 (3), 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 (6), 95–127.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5 (1), 31–56.

Anderson, R., 2007. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. 1st ed. OUP Oxford.

Ardic, O.P., Mylenko, N., and Saltane, V., 2012. Access to Finance by Small and Medium Enterprises: a Cross-Country Analysis with A New Data Set. *Pacific Economic Review*, 17 (4), 491–513.

Balcaen, S., Manigart, S., Buyze, J., and Ooghe, H., 2012. Firm Exit After Distress: Differentiating Between Bankruptcy, Voluntary Liquidation and M&A. *Small Business Economics*, 39 (4), 949–975.

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 (4), 1260–1288.

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

Beneish, M.D. and Gardner, J.C., 1995. Information costs and liquidity effects from changes in the Dow Jones Industrial Average list. *Journal of Financial and Quantitative Analysis*, 30 (1), 135–157.

Cleves, M.A., Gould, W.W., and Gutierrez, R.G., 2010. *An introduction to survival analysis using Stata*. 3rd ed. Stata Corp.

Dhillon, U. and Johnson, H., 1991. Changes in the Standard and Poor's 500 List. *Journal of Business*, 64 (1), 75–85.

Florackis, C., Gregoriou, A., and Kostakis, A., 2011. Trading frequency and asset pricing on the London Stock Exchange: Evidence from a new price impact ratio. *Journal of Banking & Finance*, 35 (12), 3335–3350.

Gao, X., Ritter, J.R., and Zhu, Z., 2013. Where have all the IPOs gone? *Journal of Financial and Quantitative Analysis*, 48 (06), 1663–1692.

Gregoriou, A. and Ioannidis, C., 2006. Information costs and liquidity effects from changes in the FTSE 100 list. *The European Journal of Finance*, 12 (4), 347–360.

Grilli, L., 2005. The random-effects proportional hazards model with grouped survival data: a comparison between the grouped continuous and continuation ratio versions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168 (1), 83–94.

Gupta, J., Gregoriou, A., and Healy, J., 2014. Forecasting Bankruptcy for SMEs Using Hazard Function: To What Extent Does Size Matter? *Review of Quantitative Finance and Accounting*, Forthcoming.

Gupta, J., Gregoriou, A., and Healy, J., 2015. *Using Hazard Models Correctly: A Comparison Employing Different Definitions of SMEs Financial Distress*. No. Working Paper.

Gupta, J., Wilson, N., Gregoriou, A., and Healy, J., 2014a. The value of operating cash flow in modelling credit risk for SMEs. *Applied Financial Economics*, 24 (9), 649–660.

Gupta, J., Wilson, N., Gregoriou, A., and Healy, J., 2014b. The Effect of Internationalization on Modelling Credit Risk for SMEs: Evidence from UK Market. *Journal of International Financial Institutions, Markets & Money*, 31, 397–413.

Hegde, S.P. and McDermott, J.B., 2003. The liquidity effects of revisions to the S&P 500 index: An empirical analysis. *Journal of Financial Markets*, 6 (3), 413–459.

Horne, J.C. Van, 1970. New listings and their price behavior. *The Journal of finance*, 25 (4), 783–794.

Hosmer Jr, D.W., Lemeshow, S., and Sturdivant, R.X., 2013. *Applied Logistic Regression*. 3rd ed. New Jersy: John Wiley & Sons.

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

Keasey, K., Pindado, J., and Rodrigues, L., 2014. The determinants of the costs of financial distress in SMEs. *International Small Business Journal*, 0266242614529317.

Kim, J., 1999. The relaxation of financing constraints by the initial public offering of small manufacturing firms. *Small Business Economics*, 12 (3), 191–202.

Lin, S.M., Ansell, J., and Andreeva, G., 2012. Predicting default of a small business using different definitions of financial distress. *Journal of the Operational Research Society*, 63 (4), 539–548.

Long, J.S. and Freese, J., 2014. *Regression models for categorical dependent variables using Stata*. 3rd ed. Texas: Stata press.

Mills, M., 2011. *Introducing survival and event history analysis*. London: SAGE Publications.

Rabe-Hesketh, S. and Skrondal, A., 2012. *Multilevel and longitudinal modeling using Stata,Volume II: Categorical Responses, Counts, and Survival*. 3rd ed. Texas: Stata Corp.

Shleifer, A., 1986. Do demand curves for stocks slope down? *The Journal of Finance*, 41 (3), 579–590.

Sueyoshi, G.T., 1995. A class of binary response models for grouped duration data. *Journal of Applied Econometrics*, 10 (4), 411–431.

Table and Figures

Table 1: Sample Description

Year	Listed SMEs				Unlisted SMEs			
	Distressed	Censored	Total	% Distressed	Distressed	Censored	Total	% Distressed
1982	2	78	80	2.50	212	1,092	1,304	16.26
1983	7	82	89	7.87	238	1,095	1,333	17.85
1984	7	129	136	5.15	251	1,130	1,381	18.18
1985	13	201	214	6.07	288	1,076	1,364	21.11
1986	16	285	301	5.32	338	1,053	1,391	24.30
1987	14	289	303	4.62	352	1,119	1,471	23.93
1988	11	290	301	3.65	338	1,190	1,528	22.12
1989	13	256	269	4.83	372	1,076	1,448	25.69
1990	9	226	235	3.83	346	1,054	1,400	24.71
1991	13	214	227	5.73	325	1,034	1,359	23.91
1992	9	293	302	2.98	295	1,027	1,322	22.31
1993	13	417	430	3.02	265	1,094	1,359	19.50
1994	19	454	473	4.02	243	1,126	1,369	17.75
1995	26	462	488	5.33	236	1,122	1,358	17.38
1996	27	493	520	5.19	270	1,118	1,388	19.45
1997	20	530	550	3.64	291	1,235	1,526	19.07
1998	42	512	554	7.58	321	1,136	1,457	22.03
1999	57	436	493	11.56	371	1,023	1,394	26.61
2000	44	446	490	8.98	328	1,043	1,371	23.92
2001	33	476	509	6.48	383	993	1,376	27.83
2002	42	444	486	8.64	509	820	1,329	38.30
2003	60	355	415	14.46	459	751	1,210	37.93
2004	48	387	435	11.03	406	765	1,171	34.67
2005	37	391	428	8.64	331	772	1,103	30.01
2006	42	369	411	10.22	330	668	998	33.07
2007	52	376	428	12.15	298	624	922	32.32
2008	52	319	371	14.02	262	504	766	34.20
2009	72	279	351	20.51	302	421	723	41.77
2010	59	289	348	16.95	267	374	641	41.65
2011	32	314	346	9.25	209	354	563	37.12
2012	41	301	342	11.99	310	204	514	39.69
2013	53	312	365	14.52	216	275	491	43.99
2014	6	23	29	20.69	17	33	50	34.00
Total	991	10,728	11,719		9,873	28,486	38,359	

Notes: This table presents year-wise distribution of samples of listed and unlisted SMEs used for this study. Columns two and six report the number of firms which has experienced the financial distress event, while columns three and seven report the number of censored observation for respective years. Column four and eight show total number of firms-year observations, while columns five and nine show the percentage of distressed firms in respective time periods.

Table 2: List of Explanatory Variable

Variable	Definition	Compustat Data Item
EBITDATA	Earnings before interest taxes depreciation and amortization/total assets	EBITDA/AT
STDEBV	Short term debt/equity book value	DLC/SEQ
CTA	Cash and short-term investments/total assets	CHE/AT
RETA	Retained earnings/total assets	RE/AT
CETL	Capital employed/total liabilities	(AT – LCT)/LT
TLTA	Total liabilities/total assets	LT/AT
CAG	Capital growth; calculated as ($Capital_t$ / $Capital_{t-1}$) - 1	(AT - LCT)
TTA	Taxes/total assets	TXT/AT
LCR	ln(current assets/current liabilities)	ln(ACT/LCT)
TCTA	Trade creditors/total assets	AP/AT
FETA	Financial Expense/total assets	XINT/AT
RtoV	Absolute Returns divided by Trading Volume	----
RtoTR	Absolute Returns divided by Turnover	----

Notes: This table lists the set of covariates along with their respective definition that we use for our empirical analysis. The last column presents the specific Compustat database items that we use to estimate the covariates.

Table 3: Descriptive Statistics

Variable	Status Indicator	Listed SMEs			Unlisted SMEs		
		Mean	Median	SD	Mean	Median	SD
EBITDATA	0	-0.0590	0.0390	0.3204	-0.1610	0.0247	0.5544
	1	-0.5071	-0.3325	0.5534	-0.7688	-0.3492	0.9020
STDEBV	0	0.0715	0.0076	0.1979	0.1607	0.0348	0.4271
	1	0.1489	0.0205	0.4203	0.0419	0.0000	0.6725
CTA	0	0.3713	0.3186	0.2872	0.2270	0.1146	0.2564
	1	0.4073	0.3815	0.2928	0.1692	0.0639	0.2327
RETA	0	-1.1893	-0.2157	3.0453	-2.6794	-0.3953	6.3650
	1	-4.5096	-2.5186	6.0285	-8.6159	-3.1280	10.7868
CETL	0	4.4521	3.0671	3.7312	3.0706	1.6158	3.7060
	1	1.3229	1.0330	1.4607	0.5574	0.3444	1.4324
TLTA	0	0.3160	0.2655	0.2292	0.5870	0.4398	0.6467
	1	0.7437	0.6388	0.4626	1.4944	0.9658	1.1523
CAG	0	0.5102	0.1278	1.1174	0.3267	0.0434	1.1700
	1	-0.2107	-0.2547	0.5934	-0.1358	-0.2908	1.0686
TTA	0	0.0148	0.0003	0.0286	0.0102	0.0000	0.0251
	1	-0.0008	0.0000	0.0128	-0.0007	0.0000	0.0117
LCR	0	1.2961	1.3067	0.8158	0.6297	0.6598	1.1095
	1	0.6979	0.6361	0.8894	-0.5000	-0.3397	1.1560
TCTA	0	0.0624	0.0458	0.0597	0.1263	0.0831	0.1377
	1	0.0959	0.0662	0.0999	0.2600	0.1754	0.2226
FETA	0	0.0131	0.0048	0.0238	0.0354	0.0180	0.0542
	1	0.0357	0.0216	0.0476	0.0923	0.0548	0.0940
RtoV	0	0.1011	0.0331	0.2859	----	----	----
	1	0.1452	0.0412	0.3214	----	----	----
RtoT	0	2.5282	1.7036	3.4054	----	----	----
	1	2.8924	1.8456	3.4212	----	----	----

Notes: This table reports the mean, median and standard deviation of explanatory variables for censored and financially distressed listed and unlisted SMEs respectively. In column two, '0' represents censored group while '1' represents distressed group of firms.

Table 4: Correlation Matrix

Variable		1	2	3	4	5	6	7	8	9	10	11	12	13
EBITDATA	1	1.00												
STDEBV	2	0.24	1.00											
CTA	3	-0.18	-0.16	1.00										
RETA	4	0.74	0.26	-0.13	1.00									
CETL	5	0.17	-0.09	0.46	0.21	1.00								
TLTA	6	-0.60	-0.26	-0.16	-0.69	-0.50	1.00							
CAG	7	0.10	-0.01	0.14	0.10	0.15	-0.09	1.00						
TTA	8	0.25	-0.01	-0.02	0.16	0.05	-0.13	0.07	1.00					
LCR	9	0.39	0.02	0.50	0.45	0.68	-0.70	0.17	0.15	1.00				
TCTA	10	-0.54	-0.17	-0.18	-0.57	-0.44	0.70	-0.10	-0.10	-0.59	1.00			
FETA	11	-0.46	-0.18	-0.17	-0.52	-0.38	0.75	-0.08	-0.13	-0.56	0.49	1.00		
RtoV	12	-0.03	0.10	-0.14	-0.03	-0.10	0.13	-0.09	-0.06	-0.16	0.12	0.11	1.00	
RtoT	13	-0.02	0.04	-0.06	-0.03	-0.04	0.06	-0.05	-0.04	-0.08	0.04	0.05	0.37	1.00

Notes: This table presents correlation among the covariates analysed in this study.

Table 5: Univariate Discrete Hazard Analysis

Variable	Sign	Listed SMEs Coefficient	SE	dy/dx	R	Unlisted SMEs Coefficient	SE	dy/dx	R	Wald Sig.
EBITDATA	-	-2.9019[a]	0.1271	-5.555[a]	5	-1.3645[a]	0.0258	-21.002[a]	5	0.0000
STDEBV	+	1.1563[a]	0.1434	2.459[a]	6	-0.3341[a]	0.0280	-5.443[a]	10	0.0000
CTA	-	-0.8332[a]	0.2042	-1.531[a]	9	-1.9209[a]	0.0751	-30.121[a]	4	0.0000
RETA	-	-0.1877[a]	0.0108	-0.400[a]	11	-0.0883[a]	0.0020	-1.454[a]	11	0.0000
CETL	-	-1.0839[a]	0.0492	-0.300[a]	12	-0.8590[a]	0.0162	-7.533[a]	8	0.0000
TLTA	+	5.3174[a]	0.1984	7.750[a]	4	1.2141[a]	0.0203	19.503[a]	6	0.0000
CAG	-	-1.7616[a]	0.0941	-1.951[a]	7	-0.4608[a]	0.0142	-7.322[a]	9	0.0000
TTA	-	-34.3485[a]	2.6241	-72.691[a]	1	-32.6033[a]	0.9884	-500.719[a]	1	0.5237
LCR	-	-1.3279[a]	0.0680	-1.566[a]	8	-1.0472[a]	0.0172	-14.829[a]	7	0.0000
TCTA	+	7.8282[a]	0.6179	14.837[a]	3	4.5560[a]	0.0933	72.241[a]	3	0.0000
FETA	+	18.8655[a]	1.2655	40.608[a]	2	10.3448[a]	0.2165	168.583[a]	2	0.0000
RtoV	+	0.6934[a]	0.1356	1.396[c]	10	---	---	---		---
RtoT	+	0.0245[b]	0.0113	0.05[b]	13	---	---	---		---

Notes: a (b) [c] significant at the 1 % (5 %) [10 %] level (two-sided test). This table reports results obtained from univariate regression analysis of respective covariates for listed and unlisted SMEs respectively. Column two presents the expected sign of the coefficients, while columns three and six report the estimated coefficients of respective groups. In columns four and seven, 'SE' represents standard error of the respective estimated coefficients. 'dy/dx' is the Average Marginal Effects (AME) in percentage, reported in columns five and nine for listed and unlisted SMEs respectively. 'R' in columns six and ten show the rank of the covariates in decreasing order of the absolute value of their respective AME. The last column reports the p-values obtained from Wald test, that we use to compare the regression coefficients (to see if the coefficients are statistically different in both the groups) of listed and unlisted group of SMEs.

Table 6: Multivariate Hazard Models

Variable	Expected Sign	Listed SMEs Coefficient	SE	dy/dx	Unlisted SMEs Coefficient	SE	dy/dx	Wald Sig.
EBITDATA	-	-2.4171[a]	0.1732	-1.959[a]	-0.6768[a]	0.0320	-6.791[a]	0.0000[a]
STDEBV	+	0.9395[a]	0.1570	0.761[a]	---	---	---	---
CTA	-	-0.8512[a]	0.2653	-0.690[a]	-0.7654[a]	0.0963	-7.680[a]	0.7662
RETA	-	---	---	---	---	---	---	---
CETL	-	---	---	---	-0.4936[a]	0.0179	-4.953[a]	---
TLTA	+	---	---	---	---	---	---	---
CAG	-	-1.0708[a]	0.0901	-0.868[a]	-0.3020[a]	0.0152	-3.031[a]	0.0000[a]
TTA	-	-24.9889[a]	3.2041	-20.258[a]	-28.5456[a]	1.1268	-286.43[a]	0.1740
LCR	-	---	---	---	-0.1382[a]	0.0253	-1.387[a]	---
TCTA	+	0.7488	0.8139	0.603	---	---	---	---
FETA	+	12.0645[a]	1.4418	9.780[a]	1.3030[a]	0.2699	13.075[a]	0.0000[a]
RtoV	+	0.3113[b]	0.1374	0.252[b]	---	---	---	---
RtoT	+	0.0247[c]	0.0139	0.020[c]	---	---	---	---
Micro		-0.8707[a]	0.1962	-0.705[a]	0.7971[a]	0.0551	7.999[a]	---
Small		0.0688	0.1330	0.055	0.4492[a]	0.0478	4.507[a]	---
Age Dummies		---	---	---	---	---	---	---
Year Dummies		---	---	---	---	---	---	---
Goodness of Fit		Value	p-value		Value	p-value		
Wald chi2		731.10	0.0000		4841.52	0.0000		
Log likelihood		-2299.73			-15319.33			
AUROC								
Within Sample		0.8735			0.8563			
Holdout Sample		0.8172			0.8518			
Number of observations		11,719			38,359			
Distressed		991			9,873			
Censored		10,728			28,486			

Notes: a (b) [c] significant at the 1 % (5 %) [10 %] level (two-sided test). This table reports results obtained from multivariate regression analysis of listed and unlisted group of SMEs. Column two presents the expected sign of the coefficients. 'SE' represents standard error of the respective estimated coefficients, while 'dy/dx' is the Average Marginal Effects (AME) in percentage. The last column reports the p-values obtained from Wald test of equality of coefficients of unlisted and listed groups.

Figure 1: Hazard Curves

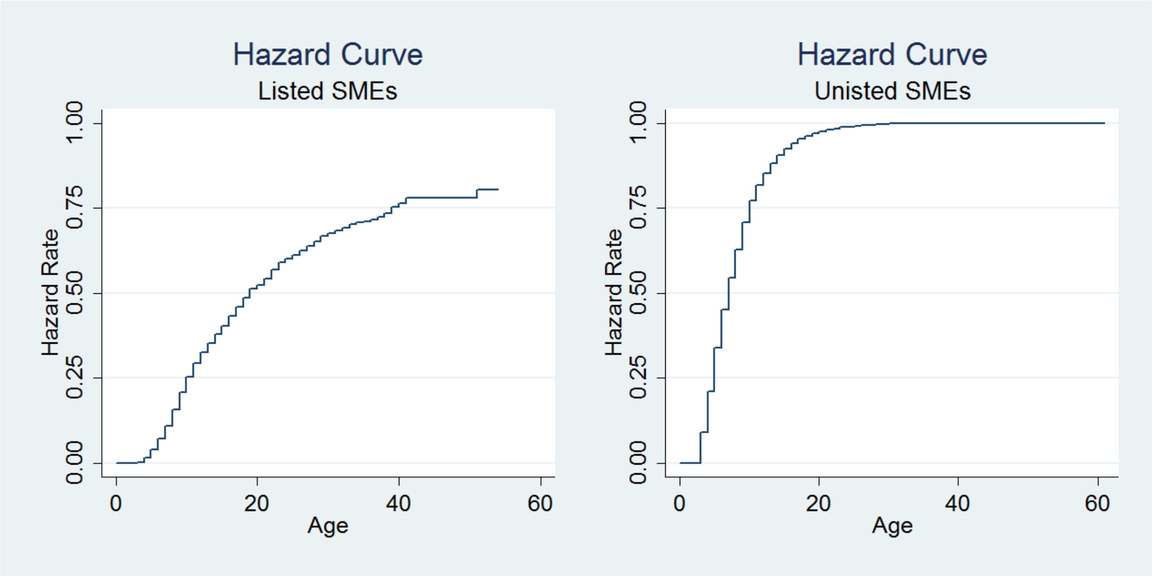

Figure 2: Area under ROC Curves