MODELLING SMALL BUSINESS FAILURES IN MALAYSIA

Nur Adiana Hiau Abdullah[1]*, Abd. Halim Hamilton Ahmad[2], Rohani Md. Rus[3], Nasruddin Zainudin[4]

[1]Prof. Dr., Universiti Utara Malaysia, MALAYSIA, diana897@uum.edu.my

[2]Mr., Universiti Utara Malaysia, MALAYSIA, abd.halim@uum.edu.my

[3]Assoc. Prof. Dr., Universiti Utara Malaysia, MALAYSIA, rohani@uum.edu.my

[4]Assoc. Prof. Dr., Universiti Utara Malaysia, MALAYSIA, nash@uum.edu.my

*Corresponding author

Abstract

Small and medium-sized enterprises (SMEs) are significant contributors to development and growth in an economy. Since SME failure is common, this study develops a failure prediction model for the small and medium-sized enterprises. By using 132 distressed and non-distressed SMEs in Malaysia during the period of 2000 through 2010 based on the logistic regression, the results for Model 1 illustrate that higher gearing and lower profitability are associated with a higher failure risk. In addition, the incorporation of the firm's age as an additional factor in Model 2 significantly improves the model's predictive accuracy. A high area under the receiver operating curve indicates that Model 2 is excellent in discriminating between the distressed and non-distressed SMEs. The Hanley and McNeil test statistic shows that both models could predict failure better than a random model. The overall prediction accuracy rate ranges from 50% to 83% and 75% to 89% for the respective Model 1 and 2 when applied to the 1-year, 2-year, 3-year and 4-year prior to distress holdout samples. Our result indicates that young SMEs rely heavily on debt and they are less efficient which led them into distressed situation. The model can detect failure as early as four years prior to the event. The developed model could be used as a refined tool to avoid possible adverse situations among the small medium-sized enterprises, creditors and investors. To the SMEs, the model could assist in early detection of distress situation; whereas to credit providers the predictors could be included in the score card for better credit decision making. Investors could potentially use the model to assess the financial well-being of companies to safeguard their interests.

Keywords: Small-medium size enterprises, business failures, emerging market, receiver operating characteristic (ROC) curve, logistic regression.

1 INTRODUCTION

Over the past decade, numerous corporate failure predicting models have been proposed and developed (see Balcaen&Ooghe, 2006, for an overview). These models, however, have focused on large firms, as financial data is more easily accessible. As the role of the small and medium-sized enterprises (SMEs) sector in the economy has become more significant recently, research on SMEs has drawn more interest from academics, practitioners and policy-makers.

In Malaysia, SMEs[1] account for approximately 99% of all business establishments and provide 56% of the country's employment. The sector constitutes 31% of the Gross Domestic Product (GDP), including 19% of the nation's total exports. Additionally, from 2004 to 2009, the value-added growth of SMEs consistently outperformed that of the overall economy, with an average annual rate of 6.3%, compared to 4.5% for the overall GDP growth (National SME Development Council, 2011). Despite their pivotal role in the economy,

[1] According to the National SME Development Council, a company is considered an SME if its annual sales turnover does not exceed RM25 million or if its workforce does not exceed 150 employees (SME Bank, 2006).

ISBN: 978-605-64453-2-3

the estimated failure rate for SMEs is 60% (Ahmad &Seet, 2009).

Presently, there is a dearth of evidence in the empirical literature on Malaysian SME failures.[2] Previous studies on Malaysian business failures have focused on listed companies (Ong *et al*., 2011). SMEs are riskier than large firms; thus, the replication of failure prediction models developed for large firms could lead to deterioration in the model's performance (Altman &Sabato, 2007). Furthermore, Agarwal and Taffler (2008) argue against direct cross-border model replication, due to reporting and insolvency code differences.

In this study, an empirical failure prediction model using Malaysian SME data is developed. This empirical model can be used as a mechanism to detect the early indication of firm failure and aid in abetter understanding of the factors that can lead to default among SMEs. Focus is placed on examining and formulating a business failure prediction model. This will be accomplished by using the Receiver Operating Characteristic (ROC) curve, an endeavor rarely considered in previous SME failure studies.

2 DATA AND METHODS

A list of distressed SMEs[3] was obtained from the CCM database[4] by selecting firms classified under a winding-up order during the period of 2000 through 2010. For each distressed company, a non-distressed match was identified based on the closest asset size. After employing data-cleaning techniques, the final sample was 132. Data for the previous two years were used in the estimation analysis because most failed companies did not submit financial reports as they approached the winding-up period, which led to a very small sample size for the year prior to failure. Potential variables were identified from previous empirical studies on firm failures. These included firm liquidity, leverage, efficiency and profitability. The model also incorporated a non-financial variable (firm age) to improve prediction accuracy (Lugovskaya*,* 2010; Pederzoli& Torricelli, 2010).

3 EMPIRICAL RESULTS

Table 1 presents the summary statistics of the variables and shows that failed firms are less profitable, have less liquidity, are highly leveraged, and are younger. Owners of distressed companies are losing money, as is reflected in the negative returns. There is an insignificant difference in the size of the distressed and non-distressed SMEs.

Table 1. Dataset properties

Variables	Non-default	Default sample	*t*-stat
Firm age (in years)	19.91	14.48	4.67**
Size (total assets)	15,193,250	14,218,700	0.27
Total liabilities to total assets	0.84	1.70	3.32**
Short-term liabilities to total liabilities	0.87	0.93	1.94*
Current assets to current liabilities	2.21	1.30	1.16
Sales to total assets	1.02	1.15	0.75
Earnings before interest and tax to total assets	0.02	-0.24	3.76**
Return on equity	0.24	-0.87	2.85**

Notes:*, ** denote significance at 10% and 5%, respectively.

[2]The scarce empirical evidence on small business finance research, especially in Malaysia, is due to the difficulty in accessing data. Under Section 165 of the Companies Act, companies are required to submit audited financial statements to the Companies Commission of Malaysia (CCM). Before the year 2000, most of the data were only available in the headoffice of the CCM. It was rather difficult to request financial statements, as the headoffice was also handling other matters pertaining to company registration. From 2000 onwards, CCM developed a database to record the financial data of registered companies in Malaysia. This system is made available to most CCM branches throughout Malaysia.

[3]The definition used to represent distressed companies was based on Part X, Section 218 of 1 (e) and (2) of the Companies Act 1965 (Revised 1973: Incorporating latest amendment-Act A1299/2007).

[4] CCM is the Companies Commission of Malaysia, an autonomous body that functions as a one-stop centre for corporate information, regulation and the development of a conducive environment. The CCM database contains financial statements and corporate information for Malaysian companies.

The stepwise logistic regression results[5] (Table 2) illustrate the debt ratio and profitability being significant for both models. Higher gearing and lower profitability should entail a higher probability of failure. This is consistent with Luppi*et al.'s* (2007) Italy findings. When non-financial variables are added to the specification, firm age is significant (Table 2: Panel B). The negative sign illustrates that young firms have a higher probability of failing, as compared to mature firms. This concurs with Shane (1996), who found that young firms have a higher probability of failure than established firms. This is because the former have a higher tendency to lose money, which could, in turn, affect their operational growth. In addition, young SMEs in Malaysia were found to rely more on debt, because they might not have adequate personal resources to fulfill their obligations.

Table 2. Stepwise logistic regression

Panel A: Model developed without non-financial variables (Model 1)

Log(PD/1-PD) =	-	0.700**	
	+	0.409**	Total liabilities to total assets
	-	4.162**	Earnings before interest and tax to total assets
Hosmer-Lemeshow test			7.635 (8 degree of freedom, *p*-value =0.470)
-2 Log likelihood			152.782
X^2 *p*-value			30.209(2)**
Cox-Snell's R^2			0.205
Nagelkerke's R^2			0.273
Area under ROC curve (AUC)			0.777 (SE=0.04, z=19.18)
Accuracy ratio			0.554

Panel B: Model developed with non-financial variables (Model 2)

Log(PD/1-PD) =	+	10.422**	
	+	0.389**	Total liabilities to total assets
	-	7.733**	Earnings before interest and tax to total assets
	-	4.082**	Ln firm age (in years)
Hosmer-Lemeshow test			7.366 (8 degree of freedom, *p*-value = 0.498)
-2 Log likelihood			106.165
X^2 *p*-value			76.825(3)**
Cox-Snell's R^2			0.441
Nagelkerke's R^2			0.588
Area under ROC curve (AUC)			0.893 (SE=0.03, z=30.95)
Accuracy ratio			0.786

Notes: Accuracy ratio = 2*(AUC-0.5), ** denotes significance at 5%. SE refers to standard errors and z refers to the z-statistic for the difference between AUC and the random model (AUC of 0.50). The Z-statistic for difference in AUCs of Model 1 and Model 2 is 2.33.

Panel B of Table 2 presents the model fit measures. The Hosmer-Lemeshow test suggests that both models are adequate and that the models fit the data (Hosmer *et al.*, p. 153, 2013). Furthermore, Cox-Snell's R-squared and Nagelkerke's R-squared tests suggest a relative increase in the model's performance when firm age is added to the specification.

A more appropriate approach to validating the predictive accuracy of the models' estimates using a

[5] For comparison purposes, an estimation using stepwise discriminant analysis also provides a similar result. It is available upon request.

logistic regression is to examine the area under the ROC curve (Bauer & Agarwal, 2014). Figure 1 presents the ROC curves for Models 1 and 2, illustrating that Model 2 has a larger area under the ROC curve than Model 1. This signifies that Model 2 is superior in its performance. Sobehart and Keenan (2001) reaffirmed this, when they explained that the area under the ROC curve (AUC) could indicate the predictive accuracy of a model.

A considerable increase in AUC is observed when the non-financial variable is added (from an AUC of 0.777 to an AUC equals 0.893). According to Hosmer*et al*. (p. 177, 2013), if the area under the ROC curve is between 0.8 and 0.9, the model is excellent in discriminating between those subjects; in this case, it is between the failed and healthy firms.

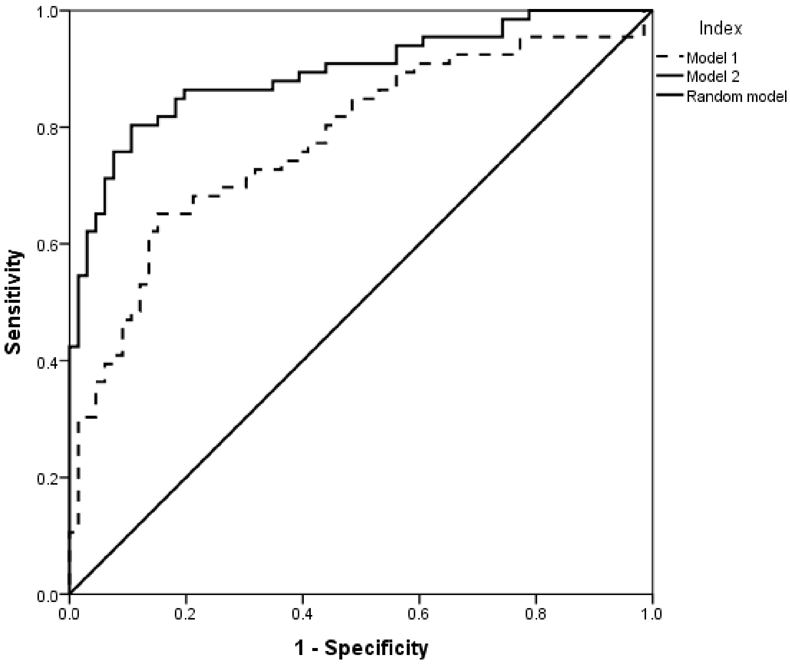

Fig. 1. Comparison of ROC curves between Models 1 & 2

To demonstrate that both models could predict failure better than a random model, the Hanley and McNeil (1982) test statistic, given by $Z = \frac{A}{SE(A)}$, is used. The standard error is:

$$SE(A) = \sqrt{\frac{A(1-A) + (n_F - 1)(Q_1 - A^2) + (n_{NF} - 1)(Q_2 - A^2)}{n_F n_{NF}}} \quad (1)$$

where: A is the area under the ROC curve, n_F is the number of failed firms and $n_{NF}$is the number of non-failed firms. Q_1 is given by $\frac{A}{2-A}$ and Q_2 is given by $\frac{2A^2}{1+A}$. The z-statistic values illustrate that both models significantly outperformed the random classification model with z-statistics of 19.18 and 30.27, respectively. Furthermore, to compare the predictive accuracy of both models, the difference in the areas under the two ROC curves is evaluated using the normally distributed z-statistic proposed by Hanley and McNeil (1983):

$$z = \frac{A_1 - A_2}{\sqrt{(SE(A_1))^2 + (SE(A_2))^2 - 2rSE(A_1)SE(A_2)}} \quad (2)$$

The z-statistic of 2.33 signifies that the inclusion of the non-financial variable in Model 2's specification significantly improves the model.[6] The same result holds true for the accuracy ratio.

We then use the estimated models to examine the accuracy rate when these models are applied to the holdout sample 1-year, 2-years, 3-years and 4-years prior to the distress. Table 3 shows that the models can predict a firm's failure as much as four years in advance of the failure. The overall rate of correct predictions ranged from 50% to 83% and 75% to 89% for Models 1 and 2, respectively. Nevertheless, the overall classification accuracy rate reduced as we moved further away from the distress situation.

Table 3: Classification accuracy

<table>
<tr><th></th><th colspan="2">Default</th><th colspan="2">Non-default</th><th colspan="2">Overall</th></tr>
<tr><th colspan="7">Training sample</th></tr>
<tr><td>Model 1</td><td colspan="2">60.6%</td><td colspan="2">86.4%</td><td colspan="2">73.5%</td></tr>
<tr><td>Model 2</td><td colspan="2">78.8%</td><td colspan="2">89.4%</td><td colspan="2">84.1%</td></tr>
<tr><th colspan="7">Holdout sample</th></tr>
<tr><td></td><td>Model 1</td><td>Model 2</td><td>Model 1</td><td>Model 2</td><td>Model 1</td><td>Model 2</td></tr>
<tr><td>1-year prior sample</td><td>76.9%</td><td>80.8%</td><td>88.5%</td><td>96.2%</td><td>82.7%</td><td>88.5%</td></tr>
<tr><td>2-years prior sample</td><td>37.5%</td><td>75.0%</td><td>62.5%</td><td>75.0%</td><td>50.0%</td><td>75.0%</td></tr>
<tr><td>3-years prior sample</td><td>42.3%</td><td>73.9%</td><td>87.4%</td><td>84.7%</td><td>64.9%</td><td>79.3%</td></tr>
<tr><td>4-years prior sample</td><td>29.2%</td><td>64.6%</td><td>90.8%</td><td>88.5%</td><td>60.0%</td><td>76.6%</td></tr>
</table>

Notes: Cases correctly classified. Model 1: Model developed without non-financial variables, Model 2: Model developed with non-financial variables.

4 CONCLUSION

This study contributes to the failure prediction literature for privately-owned SMEs in Malaysia. The findings indicate that the inclusion of the age of firms, a non-financial variable, is important in predicting failure among small and medium-sized enterprises. There is a larger area under the receiver operating characteristic curve and a higher accuracy ratio once age is included in the model. The developed model could be used as a refined tool to avoid possible adverse situations among the small to medium-sized enterprises, creditors and investors. For SMEs, the model could assist in the early detection of a distress situation; whereas for credit providers, the predictors could be included in the scorecard for better credit decision making. Investors could potentially use the model to assess the financial well-being of companies to safeguard their interests.

5 ACKNOWLEDGEMENT

This research was financially supported by the Ministry of Education, Malaysia under the Fundamental Research Grant Scheme.

REFERENCE LIST

Agarwal, V. and Taffler, R.J. (2008). Comparing the performance of market-based bankruptcy prediction models. *Journal of Banking and Finance*, **32**, 1541-51.

Ahmad, N.H. and Seet, P-S. (2009). Dissecting behaviours associated with business failure: A qualitative study of SME owners in Malaysia and Australia. *Asian Social Science*, **5**, 98-104.

Altman, E. and Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the US market. *Abacus*, **43**, 332-57.

Balcaen, S. and Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems, *British Accounting Review.* **38**, 63-93.

Bauer, J. and Agarwal, V. (2014). Are hazard models superior to traditional prediction approaches? A comprehensive test. *Journal of Banking & Finance*, **40**, 432-42.

Department of Statistics. (2011). Social Statistics Bulletin, Department of Statistics, Malaysia.

[6]Since the two models are applied to the same data, their correlation is likely to be positive. Hence, assuming a zero correlation means the denominator in Equation (2) has been over-estimated (i.e., the z-statistic is biased downwards).

Hanley, J.A. and McNeil, B.J. (1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29-36.

Hanley, J.A. and McNeil, B.J. (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases, *Radiology*, **148**, 839-43.

Hosmer, Jr., D. W., Lemeshow, S. and Sturdivant, R.X. (2013). Applied Logistic Regression 3rd Edition, New Jersey: John Wiley & Sons, Inc.

Lugovskaya, L. (2010) Predicting default of Russian SMEs on the basis of financial and non-financial variables. *Journal of Financial Services Marketing*, **14**, 301-13.

Luppi, B., Marzo, M.,and Scorcu, A.E. (2007). A credit risk model for Italian SMEs, Working Papers No. 600, DipartimentoScienzeEconomiche, Universita' di Bologna, Bologna.

National SME Development Council. (2011). SME Annual Report 2009/2010: Transformation into the New Economic Model, SME Corporation Malaysia, Kuala Lumpur.

Ong, S.W., Yap, V.C. and Khong, R.W.L. (2011). Corporate failure prediction: A study of public listed companies in Malaysia. *Managerial Finance,***37**, 553-64.

Pederzoli, C. and Torricelli, C. (2010). A parsimonious default prediction model for Italian SMEs. *Banks and Bank Systems*, **5**, 5-9.

Soberhart, J. and Keenan, S. (2001). Measuring default accurately. *Risk*, March, 31-33.

Shane, S. (1996). Hybrid organizational arrangements and their implications for firm growth and survival: A study of new franchisors. *Academy of Management Journal,*39(1), 216-234.

SME Bank. (2006). SME Bank Annual Report 2005, SME Bank, Kuala Lumpur.

ISBN: 978-605-64453-2-3