# Predicting financial distress in South Africa: the role of macroeconomic factors

By

MDUDUZI PERCY MPHO NTULI

Research report submitted in fulfilment of the requirements for the degree of Master of Management in Finance & Investment

In the

FACULTY OF COMMERCE, LAW AND MANAGEMENT

WITS BUSINESS SCHOOL

At the

UNIVERSITY OF THE WITWATERSRAND

SUPERVISOR: Dr ODONGO KODONGO

**Declaration**

I, Mduduzi Percy Mpho Ntuli declare that the research work reported in this dissertation is my own, except where otherwise indicated and acknowledged. It is submitted for the degree of Master of Management in Finance and Investment in the University of the Witwatersrand, Johannesburg. This thesis has not, either in whole or in part, been submitted for a degree or diploma to any other universities

Signed: ………………………………………………….. Date: 22 June 2018

# Abstract

Recent South African studies into corporate financial distress prediction have focused on multivariate models in which only financial ratios, computed from annual financial statements, are used as predictors. These South African financial ratio models, developed using multiple discriminant analysis, are limited to a corporate's internal financial condition without consideration of the external factors that could result in financial distress. The increase in South African corporate failures during or following economic crisis has made it clear that external factors are important predictors to be applied in conjunction with internal financial ratios. International research has shown that financial distress depends on a corporate's financial position, characteristics, and macroeconomic conditions. Different techniques have been applied by researchers to develop prediction models with these factors as predictors. Altman et al. (2017) used international data, excluding South Africa, to develop a logistic regression model based on financial ratios, firm characteristics, industrial sector and country risk. An empirical study is required to determine the applicability of this new model to predicting South African corporate failures. Furthermore, it is pertinent to determine if using South African corporate failure data improves the accuracy of the international model.

In this report, the focus is on 99 Johannesburg Stock Exchange listed firms from 2000 to 2015 of which 34 failed and 65 are non-failed. For each firm in the sample, cross-sectional data consisting of financial ratios, firm characteristics, industrial sector and macroeconomic variables is collected. The financial ratios considered are: working capital to total assets; retained earnings to total assets; earnings before interest and tax to total assets; and book value of equity to total debt. The size of the firm and its age are taken as non-financial characteristics. Industrial sectors considered are: construction & materials; industrial goods & services; and technology industrial sectors. Annual economic growth, annual inflation, and average annual lending rates are taken as indicators of the macroeconomic environment. Six hypotheses are postulated to investigate the accuracy of a logistic regression or logit model in predicting financial failure within one year. Seven different models examine the classification performance when four financial ratios are combined with firm size, firm age, macroeconomic variables, and different industrial sectors. The classification accuracy of the models is measured by the Area Under the ROC Curve (AUC).

Empirical results indicate that all models, using South African firm data, perform better than the international model and have predictive accuracies above 90% for in-sample predictions within one year. The highest predictive accuracy is achieved by a model containing all variables investigated in this study. For out-of-sample predictions, this model correctly classified 91% of non-failed and 100% of failed firms one year in advance. This study provides evidence that classification accuracy of

financial ratio models can be improved by considering firm characteristics, the industry sector in which the firm operates, and the macroeconomic environment.

# Acknowledgement

I would like to thank my mother, Sindisiwe Ntuli for her continuous moral support, encouragement and best wishes throughout my studies. Her eternal pursuit of education and knowledge against many odds has been a source of inspiration my entire life. This report is dedicated to her indomitable spirit, may it never waver.

I thank my brother, Skhulusabo Skhosana for always being there to listen and encourage me. To my family and friends, thank you for being patient, understanding and supportive of my quest for knowledge. I also thank the friends I have met during my studies who travelled this academic road with me and made it enjoyable.

Special thanks to Dr Odongo Kodongo, my supervisor, who guided me through this research and offered his expert advice unreservedly. His appraisals at every stage encouraged me to be diligent in my work and complete this work successfully.

# Table of Contents

# Chapter 1 - Introduction

## 1.1. Purpose of the study

Financial distress prediction models are tools used by corporate shareholders, stakeholders, and management to evaluate the financial condition of a corporation. Initial prediction models comprised only financial ratios, derived from annual financial statements, in univariate or multivariate form. New techniques were developed to improve the accuracy of multivariate models. Recent developments have sought to include firm characteristics and external factors to improve predictive accuracy of models. Using international corporate failure data, studies have reported higher predictive accuracies when extending financial ratios with firm characteristics and external factors. The main purpose of this study is to determine if financial ratios, firm characteristics, industrial sector and macroeconomic conditions are significant predictors of corporate failures in South Africa.

## 1.2. Context of the study

The prediction of corporate failure is an important field of study for several reasons. Firstly, an accurate failure prediction model can highlight the need for preventative or corrective actions in firms to avoid the large costs and impact on stakeholder that business failure can have. Secondly, the measurement of company performance and causes of company failure is critical for companies operating in negatively spiralling global and local economies which increases vulnerability to failure. Thirdly, opportunities for researching corporate failure prediction have increased due to the availability of data and different statistical techniques. Finally, failure prediction models are necessary to accurately assess a firm's financial situation beyond the assessment of independent auditors and management (Balcaen & Ooghe, 2006).

According to Statistics South Africa, since January 2000 there has been 55 603 companies and closed corporations that have liquidated as per Companies Act of 2008. The yearly liquidations are shown in Figure 1. From the graph, the effect of the global financial crises that resulted in a recession in South Africa between 2008 and 2009 is observed to have resulted in an increase in corporate failures. Shrinking world economic growth, which also affected South Africa, between 2000 and 2001 had a similar effect of increased levels of corporate failures.

There have been high profile companies in financial distress such as Edcon and Stuttafords, two companies in the retail sector. In the case of Edcon, a publicly traded company, macroeconomic factors such as high inflation and depreciation of the currency placed the company under pressure due to its capital structure of high debt and interest costs (Casiraghi & Monteiro, 2016). To avoid business rescue, Edcon was taken over by creditors in multibillion debt-to-equity swap. Stuttafords, a 159-year-old unlisted private company, once a top retailer, was a victim of a depressed retail sales growth environment as consumer spending slowed and competition increased. As a result, Stuttafords permanently ceased to exist on 1 August 2017 and became one of South Africa's largest corporate failures in recent history (Shevel & Rose, 2017).

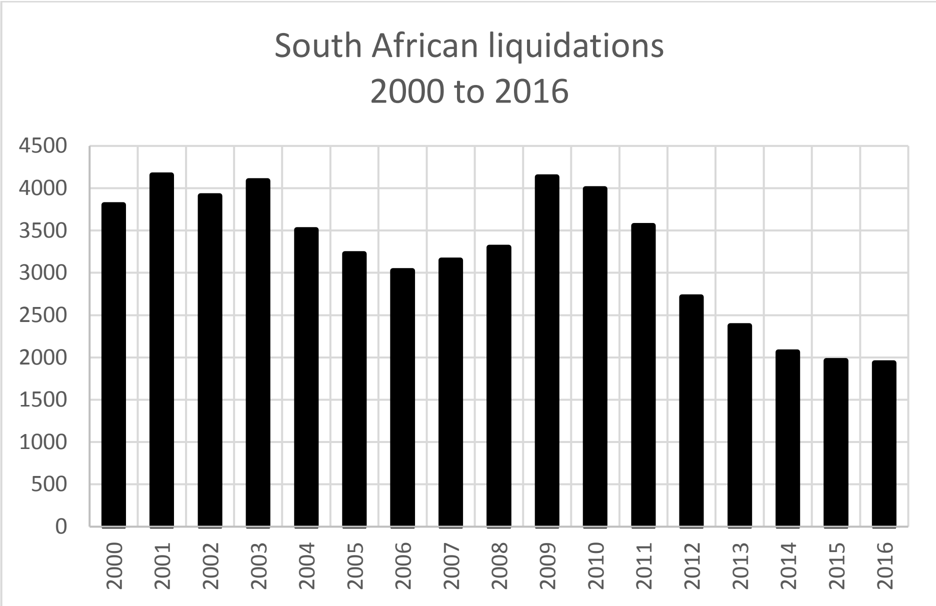


*Figure 1: South African corporate liquidations from 2000 to 2016*

The main causes of such corporate failures can be divided into: human; internal and external; structural; and financial. Human causes relate to issues surrounding leadership, management and employee skills. Internal causes of failure relate to the organisational culture and managerial strategy formulation and implementation, while external factors relate to the macroeconomic environment firms operate in. Structural causes relate to the loss of resources and knowledge capacity; however, this can also include firm age, size, and life cycle stage (Pretorius, 2008).

Many researchers have studied corporate failure to statistically model failure based on the causes identified by Pretorius (2008). The use of financial statements to determine the financial condition of a firm was pioneered by Beaver (1966) when he developed the univariate failure prediction model. In this model, financial ratios were calculated from financial statement data to determine optimal cut-offs that predict failure or continued healthy operations. This work was extended by the seminal study by Altman (1968) in which he developed a multivariate discriminant model. Using five financial ratios as independent variables, Altman created the z-score model for manufacturing firms. Altman extended this work by creating a five-variable z'-score model for private firms. Finally, Altman proposed a z"-score model using four-variables to predict failure of private non-manufacturing firms. Other researchers studied failure prediction, including the seminal work of Ohlson (1980) that proposed a novel logistic regression model. With over 50 years of failure prediction studies, the choice of financial ratios as independent variables in models still requires a theoretical framework, however, the is no general theory that describes the phenomenon. As such, there is no consensus on which financial ratios lead to the most accurate failure prediction.

Balcaen and Ooghe (2006) found that many studies in literature have found that firm size has an influence on the likelihood of failure. Small firms are more likely to fail than large firms. Large firms are thought to have a lower probability of failure due to advantages earned through experience and influence over stakeholders. Empirical results in literature provide evidence that the macroeconomic environment is a significant explanatory factor for corporate failure. This is evident during economic downturns where the corporate failure rate is higher.

## 1.3. Problem statement

This study focuses on the financial condition, structural, and external causes of firm failure. The financial condition is determined from annually published financial statements which consist of balance sheet, income statement, and cash flow information over an operating cycle. Structurally, the size and age of a firm have been found, in literature, to be related to risk of failure. Smaller firms are exposed to capital market discrimination and are likely to have inexperienced management, both of which increase the likelihood of failure according to conventional wisdom (Pretorius, 2008). Risk of failure is higher in the first few years of a firm's life, increases to a peak point, and declines as the firm ages (Pretorius, 2008). The external factors are proxied by the macroeconomic environment forces that could affect the financial condition of a firm as well as different industry sectors.

In a recent study, Altman, Iwanicz-Drozdowska, Laitinen, and Suvas (2017) developed a logistic regression model taking into account all three causes of failure in an international context. Macroeconomic conditions were proxied by individual country risk as measured by Standard & Poor's Risk Rating. Country-specific macroeconomic factors were not studied.

This study analyses the prediction performance of the Altman et al. (2017) financial ratio only model using South African sample data. The applicability of their model to South Africa is validated by determination of its accuracy to discriminate failed and non-failed publicly listed firms. This study goes further by investigating the impact of firm size, firm age, and three macroeconomic variables, and industry sectors on model accuracy. Understanding the significance of each firm characteristic, industrial sector in which a firm operates, and macroeconomic environment on corporate failure will offer all stakeholders a means of forecasting the financial condition of a firm in one year.

## 1.4. Research objectives

The primary objective of the research is to determine the applicability of failure prediction models from developed economies on firms operating in South Africa's developing economy. To achieve this, the financial ratio only logistic regression model developed by Altman et al. (2017) will be applied to South African data, consisting of failed and non-failed firms, and its predictive accuracy determined.

The secondary objective of the research is to develop a failure prediction model using logistic regression analysis on South African data. A base model will be developed by re-estimating the coefficients of the financial ratio only logistic regression model of (Altman et al., 2017) and determining its predictive accuracy. Different models will be developed to determine the effect on predictive accuracy of this financial ratio only model when firm size, firm age, macroeconomic environment and industrial sector are explicitly accounted for.

## 1.5. Significance of the study

Most failure prediction models in literature which use financial ratios as predictors are focused on developed economies. Empirical investigations have extended such models to include firm characteristics and macroeconomic variables. Research findings indicate that models applicable to developed economies are not directly applicable to developing economies such as South Africa.

The use of logistic regression analysis on South African failed firms will offer an empirical analysis of a recently developed international model by a renowned author in the field of failure prediction. Results of this analysis will give further insight into the applicability of developed economy models in South Africa. Models that investigate the impact of firm size, firm age, industry sectors and macroeconomic variables in combination with financial ratios will provide a comprehensive indication of which factors are more significant in failure prediction.

## 1.6. Assumptions of the study

The financial condition of a firm is a continuum from healthy, decline, and finally failure (Pretorius, 2009). A firm that is in decline experiences financial distress which threatens continued operations unless remedied. The South African Companies Act, 2008 no. 71 defines a company financially distressed if it cannot meet its debt obligations when they are due and payable, or it will become insolvent. Failure, in most studies, is taken as the legal definition which is court declared bankruptcy that results in immediate termination of operations. In South Africa, the Companies Act, 2008 no. 71 makes no reference to bankruptcy, but rather defines liquidation by court order. This study adopts the following definition of failure:

- Companies that have involuntarily liquidated according to the Companies Act, 2008 no. 71
- Companies that applied for voluntary, strategic, or accidental liquidation are excluded

It is assumed that all liquidated companies failed because of financial difficulties. In relation to financial statements, it is assumed that those published are an accurate and true reflection of a company's financial position several years prior to failure. Where a firm's financial statement has missing values in a specific year being considered, that firm financial year will be excluded from the sample data.

## 1.7. Delimitations of the study

The firms and industries investigated in this report must have similar financial characteristics. The following types of companies will be excluded:

- Financial institutions
- Real estate development companies
- Public sector

Financial institutions such as insurance companies, investment companies, and banks characteristically have balance sheets with higher leverage than typical firms (Wang & Dwyer, 2010). These firms are structurally different due to regulatory and capital requirements imposed on them.

The financial health of real estate development and investment companies, such as real estate investment trusts (REIT), often pivot on one or few developments. This is reflected in their financial statements which provide only a partial description of the active projects undertaken by such companies. (Wang & Dwyer, 2010)

Public sector companies run by government have a different failure risk profile due to government's unwillingness to allow them to fail. Such companies when facing financial difficulty may receive a bailout from government (Wang & Dwyer, 2010). As a result, the financial statements of public sector companies are not comparable to non-government companies.

## 1.8. Report outline

This report is structured as follows:

- Chapter 2 provides a detailed definition of failure and a review of corporate failure literature. Failure prediction models are divided into theoretical and empirical studies followed by a review of studies that included macroeconomic variables. Finally, a review of South African failure prediction studies is presented.
- Chapter 3 presents an overview of the methodology utilized in this study. The data source and data used is discussed with a detailed description of the econometric methodology applied.
- Chapter 4 presents the analysis results with thorough discussions of the findings.
- Chapter 5 concludes and makes recommendations based on the results.

# Chapter 2 - Literature Review

## 2.1. Introduction

This section provides a review of literature studies related to corporate failures. Firstly, the definitions of failure used in most studies is presented. Secondly, the theoretical and empirical studies which developed failure prediction models based on financial ratios, firm characteristics, and macroeconomic variables are discussed. Finally, South African studies into corporate failure prediction are presented.

## 2.2. Corporate failure definition

Corporate distress, also known as corporate failure, is an important field of study in literature for researchers to develop models that can predict with accuracy the onsite of the condition. The aim of corporate failure prediction studies is to discriminate failing firms from non-failing or healthy firms. Such a distinction depends entirely on the definition of "corporate failure" applied by a researcher. Corporate failure is a continuum that ranges from a financially healthy firm to one that is legally declared bankrupt with subsequent liquidation. Between the extremities of the continuum, corporate failure can present in different financial conditions such as distress, insolvency, default or bankruptcy (Altman & Hotchkiss, 2006). For research purposes, it is beneficial to define each financial condition clearly as this will affect the data sample collected for the study.

Pretorius (2009) conducted a comprehensive review that examined definitions of failure used in literature. The author proposed a universal definition of failure accounting for the multidimensional nature of the phenomenon. The proposed definition is that a healthy firm will experience a decline which is a precursor to failure. A firm is in decline when its performance deteriorates over consecutive periods or experiences distress in continuing operations. A firm in decline can be corrected by interventions that return the firm to a healthy state. If the causes of decline are not corrected, the firm will slide towards a point of failure. The author defined failure as the discontinuance or bankruptcy at which point all operations cease and judicial proceedings take effect. Pretorius highlighted that the proposed definition of failure excludes voluntary, strategic, and accidental bankruptcy or liquidation.

To shareholders, a declining firm is in financial distress when the realized rate of return on equity, while allowing for business risk, is significantly lower than prevalent rates on similar investments. In academic studies, financial distress is an economic condition where generated revenues are insufficient to cover costs (Altman & Hotchkiss, 2006). Such distress can manifest with the firm reporting negative net operating income for several years, suspending dividend payments or major restructuring (Balcaen & Ooghe, 2006). A distressed firm may remain in such a state for many years without failing to meet its current liabilities if such liabilities are not legally enforceable debts. The South African Companies Act, 2008 no. 71 defines a company financially distressed if it cannot meet its debt obligations when they are due and payable, or it will become insolvent within 6 months.

Corporate failure related to the cash position of a firm can be either technical or bankrupt insolvency. Technical insolvency is a temporary condition when a firm cannot meet its current obligations. Altman and Hotchkiss (2006) propose that technical insolvency be measured with net cash flows relative to current liabilities. Bankruptcy insolvency is a continual condition where a firm's total liabilities exceed a fair comprehensive valuation of its total assets.

Corporate failure arising from violation of loan or bond agreements is defined as default. A firm may be declared to be in default if loan repayments are delayed beyond 90 days or the firm is placed business rescue which involves postponement of interest or principal amounts by financial institutions (Balcaen & Ooghe, 2006).

The final state of a distressed firm is when it is declared to be bankrupt by a court of law. A bankrupt firm may commence liquidation proceedings or reorganization through business rescue. Balcaen and Ooghe (2006) find that most corporate failure studies apply this legal definition of failure which is bankruptcy. In South Africa the legal definition of failure is legislated in the Companies Act 71 of 2008 Chapter 6. The act uses the definition of liquidation, be it voluntary or involuntary, as the final legal state of failure. The act makes no reference to bankruptcy.

In the following sections, the definition of failure as used by the authors is reported as published. Some authors used bankruptcy, insolvency, or default to collect their data and perform analysis.

## 2.3. Theoretical Failure Prediction Models

### 2.3.1. Univariate Analysis

The use of financial ratios in corporate failure prediction was first proposed by Beaver (1966) in his seminal work on a univariate discriminant analysis model (Balcaen & Ooghe, 2006). In univariate analysis, multiple financial ratios were analysed individually to determine how accurately each ratio classifies a firm as failed or non-failed. The condition of a firm was viewed as a dichotomous state of failed or non-failed. Financial ratios were used to discriminate firms between these two groups, failed and non-failed (Altman, 2000). Beaver (1966) determined optimal cut-off points for each financial ratio, points at which the percentage of misclassification is minimized, to classify a firm as failing or non-failing. If a firm's financial ratio is below the cut-off point it is classified as failing, else it is classified as non-failing. For those financial ratios where higher values indicate possibility of failure, the opposite of the classification rule was applied. For one-year prior to failure, the analyses identified that the cash flow to total debt ratio correctly classified 87% of the sample firms. The same ratio correctly classified 78% five-years prior to failure. From these results, the author concluded that financial ratios can be used in the prediction of firm failure for up to five-years prior to failure.

In Beaver (1968), the initial work on univariate analysis was extended with a study based on sample data consisted of financial statements of 79 failed and non-failed firms from 1954 to 1964. The sample of non-failed firms was matched to the failed firms by industry and asset size. In total, the sample data represented 38 different industries and asset sizes ranging from $600 000 to $45 million. For one-, two- and three-years prior to failure, the cash flow to total debt and net income to total assets correctly classified 87%, 79%, and 77% of the sample data, respectively. Cash flow to total debt ratio correctly classified 78% of failures five-years prior to failure. Net income to total assets ratio correctly classified 72% of failures five-years prior to failure.

### 2.3.2. Multiple Discriminant Analysis

Using a sample of 66 publicly traded manufacturing firms, classifying 33 firms as bankrupt and 33 firms as non-bankrupt, Altman derived data from financial statements one reporting year before filing for bankruptcy (Altman, 1968). These firms were selected and categorised by industry and size. The asset size was limited to USD 1 – 25 million which eliminated very large firms which for which bankruptcy was rare at the time. Smaller firms were eliminated from the study due to lack of data. The financial data collected for firms in both groups were from the same years to allow comparability. Altman compiled a list of 22 financial ratios from the firm's financial statements to use as variables for failure prediction. These variables were classified into five categories: activity, liquidity, solvency, profitability, and leverage. Altman used multiple discriminant analysis (MDA) to formulate a single discriminant score calculated from multiple financial ratios, called the z-score, that distinguishes between bankrupt and non-bankrupt firms. The final discriminant function estimated by Altman (1968) consisted of five independent variables as follows:

$$z = 1.2x_1 + 1.4x_2 + 3.3x_3 + 0.6x_4 + 1.0x_5$$

where

$$x_1 = \frac{Working\ Capital}{Total\ Assets}$$

$$x_2 = \frac{Retained\ Earnings}{Total\ Assets}$$

$$x_3 = \frac{Earnings\ before\ Interest\ and\ Taxes\ (EBIT)}{Total\ Assets}$$

$$x_4 = \frac{Market\ Value\ of\ Equity}{Book\ Value\ of\ Total\ Liabilities}$$

$$x_5 = \frac{Sales}{Total\ Assets}$$

This resulting z-score model has discriminant coefficients with positive signs indicating that a firm with a high potential of bankruptcy will have a lower discriminant score. Altman (1968) concluded that firms with a z-score greater than 2.99 are not at risk of bankruptcy, while firms with a z-score below 1.81 are all bankrupt. Firms with z-scores between 1.81 and 2.99, defined as a "grey area", where classification between bankrupt and not-bankrupt was susceptible to error. Following further analyses, Altman proposed a critical z-score value of 2.675 to best discriminate between bankrupt and non-bankrupt firms and thus avoiding the grey area.

Altman (1968) categorised prediction errors as Type I and Type II. Type I errors occur when firms are predicted as non-bankrupt when it does file for bankruptcy. Type II errors occur when firms are predicted to go bankrupt but never file for bankruptcy or default on debt obligations. The classification errors are best illustrated using what Altman called the "accuracy matrix" shown in Table 1 below.

*Table 1: Altman (1968) accuracy matrix*

| **Actual Group** | **Predicted Group (z-score)** | |
|---|---|---|
| | *Bankrupt* | *Non-bankrupt* |
| *Bankrupt* | Correct classification | Type I |
| *Non-bankrupt* | Type II | Correct classification |

For the original data, the Type I error was found to be 6 percent, while the Type II error was lower at 3 percent within one financial reporting year prior to bankruptcy. Two years prior to bankruptcy, it was found that the Type I error was much higher at 28 percent, while the Type II error increased to 6 percent. Using data of 86 distressed companies between 1969 to 1975, 110 bankrupts from 1976 to 1995, and 120 bankrupts from 1997 to 1999, Altman and Hotchkiss re-evaluated the accuracy of the model (Altman & Hotchkiss, 2006). They found that the accuracy of the z-score model correctly predicted distressed firms between 80 and 90 per cent when using data from one financial reporting prior to bankruptcy. However, the Type II error in prediction was observed to have increased by as much as 25 percent. (Altman & Hotchkiss, 2006)

The original Altman (1968) z-score model was only valid for publicly traded manufacturing firms. This is evident in the use of the market value of equity for independent variable $x_4$ (Market Value of Equity/ Book Value of Total Liabilities). To modify the model for use in evaluating private firms, Altman substituted the book value of equity of the market value of equity in independent variable, $x_4$, and performed a complete re-estimation of the discriminant coefficients using the same data of 1968 (Altman, 2000). The resulting discriminant function, named the z'-score model, is as follows:

$$z = 0.717x_1 + 0.847x_2 + 3.107x_3 + 0.420x_4 + 0.998x_5$$

where all variables remain the same as the original model except

$$x_4 = \frac{Book\ Value\ of\ Equity}{Book\ Value\ of\ Total\ Liabilities}$$

Altman determined that firms with a z'-score above 2.90 were not at risk of bankruptcy, while firms with a z'-score below 1.23 are all bankrupt. This lower limit for bankruptcy, 1.23, compared to the original z-score lower limit of 1.80, indicates that the revised model is less reliable than the original, but only marginally less (Altman, 2000). The grey area of prediction for the z'-score model is between 1.23 and 2.90. Due to a lack of data of private financial data, the z'-score model was not tested extensively on a secondary sample of bankrupt and non-bankrupt firms. For the revised z'-score model, the Type I accuracy drops to 91 percent compared to 94 percent of the original model. The Type II accuracy remains identical at 97 percent.

To extend distress prediction to private nonmanufacturing firms, Altman in 1983 identified that the asset turnover independent variable, $x_5$, should be excluded from the model to minimise the potential industry effect (Altman, 2000). The asset turnover ratio is industry sensitive as it illustrates the sales generating ability of the firm's assets. Altman estimated a four-variable model using book value of equity for the independent variable, $x_4$. The resulting four-variable discriminant function, named the z"-score model:

$$z = 3.25 + 6.56x_1 + 3.26x_2 + 6.72x_3 + 1.05x_4$$

Firms with a z"-score below 0 are classified as distressed. The classification results of the z"-model, Type I and Type II, are identical to those of the five-variable z'-model.

Taffler (1983) applied the MDA methodology to UK manufacturing and construction firms listed on the London Stock Exchange (Agarwal & Taffler, 2007). They studied 80 selected financial ratios on a sample of 46 failed and 46 non-failed firms between 1968 and 1976. Using a stepwise methodology, Taffler determined the best discriminating financial ratios and proposed a z-score model of the form:

$$z = 3.20 + 12.18x_1 + 2.50x_2 - 10.68x_3 + 0.029x_4$$

where

$$x_1 = \frac{Profit\ before\ tax}{Current\ liabilities}$$

$$x_2 = \frac{Current\ assets}{Total\ liabilities}$$

$$x_3 = \frac{Current\ liabilities}{Total\ assets}$$

$$x_4 = \left(\frac{Quick\ assets - Current\ liabilities}{Sales - Profit\ before\ tax - depreciation}\right)365$$

Taffler determined that a firm with a positive z-score is unlikely to fail within a year, however, a negative z-score indicates risk of failure. A highly negative z-score indicates a high probability of financial failure. Agarwal and Taffler (2007) assessed the performance of this model over a period of 25 years between 1979 and 2003 in which 232 failures were reported for London Stock Exchange listed companies. Based on the one-year prior to failure financial data, the model correctly predicted that 223 (96%) of the 232 firms would fail. Agarwal and Taffler (2007) conclude that such z-score models perform well relative to other prediction models, however, caution against using models developed from country specific firm data, such as the US specific Altman z-score models, to other countries.

### 2.3.3. Conditional Probability Models

Ohlson (1980) pioneered the use of Logistic Regression Analysis (LRA), also referred to as the Logit method, in corporate failure prediction. In his work, the LRA model was used to determine the probability of a firm failing, when it belongs in some population, within a prescribed time period. Data was collected for 105 failed and 2 058 non-failed companies from 1970 to 1976. Nine independent variables, of which 6 were financial ratios, were studied to formulate three models. The first model predicted failure within one year; model two predicted failure within two years, provided the firm did not fail with the subsequent year; model three predicted failure within one or two years. The study identified four financial statement factors that are statistically significant for assessing the probability of failure: firm size; financial structure; performance measure(s); current liquidity. Firm size, an independent variable, was calculated as the logarithm of total assets divided by the GDP price-level index. Ohlson argued that this treatment of firm size assures real-time implementation of the model. Financial structure was measured using the financial ratio of total assets to total liabilities. Performance was measured using net income to total assets or cash from operations to total liabilities or a combination of both. Current liquidity was measured using working capital to total assets or current liabilities to current assets or a combination of both. The effect of current liquidity was found to not be as pronounced as the other 3 factors. To predict failure within one year, Ohlson determined the following Logit function:

$$L_{Ohlson} = -1.32 - 0.407x_1 + 6.03x_2 - 1.43x_3 + 0.0757x_4 - 2.37x_5 - 1.83x_6 + 0.285x_7 - 1.72x_8 - 0.521x_9$$

where

$$x_1 = \log\left(\frac{Total\ Assets}{GDP}\right)$$

$$x_2 = \frac{Total\ Liabilities}{Total\ Assets}$$

$$x_3 = \frac{Working\ Capital}{Total\ Assets}$$

$$x_4 = \frac{Current\ Liabilities}{Current\ Assets}$$

$$x_5 = 1\ if\ Total\ Liabilities > Total\ Assets, 0\ otherwise$$

$$x_6 = \frac{Net\ Income}{Total\ Assets}$$

$$x_7 = \frac{Cash\ Flow\ from\ Operations}{Total\ Liabilities}$$

$$x_8 = 1\ if\ Net\ Income < 0\ for\ last\ two\ years, 0\ otherwise$$

$$x_9 = \Delta Net\ Income$$

Ohlson determined a probability of failure cut-off point of 0.038 above which a firm could be classified as failed. Using this cut-off point which minimises the sum of the errors, the classification error of the LA model for the sample data was 17.4% for non-failed firms, while for failed firms it was 12.4%.

Zmijewski (1984) proposed using a Probit model to predict failure of industrial firms. In a Probit model, the dichotomous dependant variable, failure or non-failure, is assumed to have a normal cumulative distribution function (Gujarati & Porter, 2009). Using a sample of 40 bankrupt and 800 non-bankrupt industrial firms collected between 1972 to 1978, Zmijewski developed a Probit model using financial ratios that measured firm performance, leverage, and liquidity.

Grice and Dugan (2003) re-estimated the Ohlson (1980) and Zmijewski (1984) models using data collected for time period, industries, and financial conditions different from those used to develop the original models. The results of this study indicated that the accuracy of the models improved when the coefficients were re-estimated suggesting the relation between financial ratios and failure prediction changes over time.

Altman et al. (2017) conducted a study that expanded the historical work of Altman's original z-score class of models, which were based on United States firms, to an international context. They analysed the performance of the z-score model for 31 European and three non-European countries using different modifications of the original model. United States, China and Columbia were the three non-

European countries in the study. The firms in the international sample were mostly privately held with a considerable number from non-manufacturing industries. Firms in the financial sector were excluded from the study. The international sample data consisted of 38 215 failed and 2 602 563 non-failed firms from 2007 to 2010. The study applied a LRA with the four variables from Altman's z"-score model, firm characteristics, and a macroeconomic variable. Seven hypotheses were tested to determine model performance based on international and individual country data:

- Re-estimating the coefficients of the z"-score model improves its classification accuracy
- The prediction accuracy of the logistic regression version of the z"-score model is higher than that of the MDA estimated model
- Including the effect of the year of bankruptcy increases prediction accuracy
- Including the effect of firm size increases prediction accuracy
- Including the effect of firm age increases prediction accuracy
- Including the effect of industry increases prediction accuracy
- Including the effect of country risk increases prediction accuracy

Altman et al. (2017) measured model classification performance by the Area Under the Curve (AUC) measure calculated from the Receiver Operating Characteristics (ROC) curve. The AUC results showed that the original z"-score model is extremely robust across countries and over time. The re-estimation of coefficients using MDA marginally improved the classification performance. The re-estimation of the model using LRA was found to be similar to the re-estimated MDA model for most countries. They found that the use of additional variables generally improved the classification accuracy of the original model, but the results were found to be dependent on each country's distribution of failed and non-failed firms. The effects of including the year of bankruptcy and firm size were found to be the strongest on classification accuracy. When all additional variables were included in the same model, the classification performance increased significantly. Based on these results, the authors state that it is possible to create more efficient country models using the four original variables from the z"-score model with a set of additional variables. They postulate that for most countries, the classification accuracy can be improved with country specific estimation. Finally, the authors recommended introduction of macroeconomic data as a modification and extension to their study. No discussion was presented on which macroeconomic variables can be included in prediction models.

### 2.3.4. Hazard Analysis

Shumway (2001) argued that cross-sectional models, which observe each bankrupt firm's data one-year prior to bankruptcy, ignore data on healthy firms that eventually file for bankruptcy and thus introduce selection bias into the models. To resolve this shortcoming of cross-sectional models, Shumway developed a hazard model by explicitly accounting for time. The dependant variable in a hazard model is the time spent by a firm in a healthy group. The hazard statistical estimation method is also referred to as survival analysis and duration analysis in literature. A firm's risk of bankruptcy changes with time, and its health is a function of its latest financial data and its age. Shumway discussed three advantages of using hazard models over cross-sectional models to predict bankruptcy: control for each firm's period at risk by capturing that some firms file for bankruptcy after many years of being at risk while other firms fail in their first year; incorporate explanatory variables that change with time such as financial time-series data, macroeconomic variables, and firm age; produce more efficient out-of-sample forecasts by utilising more data as each year of financial data is an observation (Beaver, McNichols, & Jung-Wu, 2005). Shumway studied 300 public USA bankruptcies between 1962 and 1992 to propose a hazard model using accounting ratios and market data as independent variables. The ratio of net income to total assets, and the ratio of total liabilities to total assets, used by Zmijewski (1984), we identified as the most statistically significant. The market variables included in the model accounted for: relative firm size as the logarithm of each firm's market capitalization to the total market size; past excess return; and standard deviation of each firm's stock returns. The discrete time hazard function was estimated using logarithmic regression analysis by taking each year in which a firm survives as a sample data observation (Shumway, 2001). The model classified 75% of failing firms in the top decile of firms ranked annually by bankruptcy probability.

### 2.3.5. Option Pricing and Contingent Claims

In credit risk analysis, firms default when they fail to service debt obligations, which influences the firm's equity returns. This default risk induces lenders to require a premium on the risk-free rate of interest from borrowers. This premium is an increasing function of the probability of default of an individual firm. Vassalou and Xing (2004) proposed a purely market-based model based on Merton's (1974) option pricing model to compute monthly default measures for individual firms and assess the effect of default risk on equity returns. The proposed model estimated default likelihood indicators using the contingent claims methodology of Black and Scholes and Merton (BSM) for individual firms using market data. Following the methodology of Merton, equity of the firm is viewed as a European call option on the firm's assets, thus asserting that shareholders have residual claim after all other obligations have been met following default. The strike price of the call option is the book value of the firm's liabilities or face value of the firm's debt. When the firm's assets are worth less than the face value of debt at maturity, the value of the equity is zero. Thus, shareholders effectively sell the firm to debtholders for the face value of the debt which results in debtholders holding a portfolio of riskless debt and a short put option on the firm's assets (Reisz & Perlich, 2007). In developing the model, Vassalou and Xing observed that the volatility of a firm's assets, which is an input into the BSM formulas, provides critical information about a firm's default probability. This volatility was calculated using an iterative procedure using an initial value calculated from 12-month daily equity data.

Hillegeist, Keating, Cram, and Lundstedt (2004) proposed two modifications to the standard BSM formula in calculating default probabilities using only market data. The first modification was substituting expected market return on assets for the risk-free rate of interest. The second modification was accounting for dividends, defined in terms of total assets, as a deduction on the expected return on assets. The model was tested on a sample of 516 bankruptcies between 1979 and 1997. Results showed that the BSM model using these modifications is a powerful proxy for the probability of bankruptcy.

Bharath and Shumway (2008) proposed a simplified method for calculating the implied probability of default for a Merton (1974) option pricing default prediction model. Where Vassalou and Xing (2004) applied an estimate for the expected annual return of the firm's assets, Bharath and Shumway (2008) set the expected return on the firm's assets equal to the firm's stock return over the previous year. This "naïve" simplification allowed for past market data to be incorporated into the model and simplified calculations of default probability estimations. In the analysis, it was confirmed that the naïve probability is a significant default predictor. Using a sample of 15 018 firms and 1449 defaults, default was predicted using a hazard model with time to default as the dependent variable.

Reisz and Perlich (2007) developed a contingent claim model in which the stock price of a firm is viewed as a European down-and-out barrier option on the underlying assets of the firm. The barrier, specified by debtholders or a level of assets value below which investors will not invest in the firm, is the trigger for bankruptcy. When a barrier is reached, bankruptcy is initiated followed by debtholders taking over the firm. Such debtholders are viewed as holding a portfolio consisting of riskless debt, a short put option on the firm, and a long down-and-in call option on the firm's assets. The firm's asset value, volatility and bankruptcy barrier are implied from market equity data. Using a sample of 5 784 bankrupt industrial firms between 1998 and 2002, the down-and-out barrier model performed better than other option pricing models.

Jackson and Wood (2013) propose a naïve of simplified down-and-out call barrier option contingent claims model, similar to the Bharath and Shumway (2008) simplification of the European call option model. The barrier option framework is an extension of the European call model. In the barrier option, the underlying value of firm equity is treated as a down-and-out call option. In such a model, debt holders own a portfolio of risk-free debt and a down-and-out call option on the firm's assets which can be exercised should the value of the firm fall below a predetermined barrier. The implied barrier level specific to individual firms can be estimated using the market value firm traded equities or an appropriate proxy. Jackson and Wood (2013) simplified the barrier model by using the naïve equations of Bharath and Shumway (2008), using book value of total debt and setting the barrier to the same level as the firm's total liabilities. To further simplify the model, they assume no dividends, zero rebate, costless insolvency proceedings, and set the return on assets equal to the risk-free rate. This naïve down-and-out call model was applied for one-year prior to default and found to compare well with the European call model of Bharath and Shumway (2008). Charitou et al. (2013) studied the BSM default prediction model and extended the work of Bharath and Shumway (2008) by estimating volatility directly from market observed returns on the firm's value. Volatility was estimated using a 60-month return window based on either monthly firm value return or risk-free rate obtained from the 1-month US Treasury bill rate. Charitou et al. (2013) argued that this estimation of volatility from historic firm value return data was more parsimonious than the naïve model of Bharath and Shumway (2008). The model was tested on a sample of 1212 bankrupt US non-financial firms between 1985 and 2009. In total, including non-failed and failed firms, the data set consisted of 120 607 firm-year observations. Results showed that volatility estimated from historical firm value returns performed better that the naïve model of Bharath and Shumway (2008).

Credit risk modelling is important for commercial debt issuers to predict loan default by debt holders. Major international financial institutions have developed in-house credit risk models to minimize losses associated with loan defaults. Kollár and Gondžárová (2015) studied three credit risk models: CreditMetrics by JP Morgan; RiskCalc by Moody's; and CreditRisk+ by Credit Suisse's. The models were characterised as either default-mode or mark-to-market. CreditRisk+ and RiskCalc are default-mode type models that focus on predicting losses caused by default when considering a dichotomous state of failure or non-failure. These models are analytically based on the Merton (1974) methodology which incorporates asset value and asset volatility. CreditMetrics is a mark-to-market type which focuses on changes in loan market values and using a rating system to determine changes in the borrower's loan quality. This model is estimated using historical default rates and credit spreads.

### 2.3.6. Intelligent Techniques

Kumar and Ravi (2007) conducted a comprehensive review of bankruptcy prediction models developed using statistical and intelligent techniques between 1968 and 2005. The review categorised intelligent techniques according to the type of technique applied to solve bankruptcy prediction of banks and firms. The techniques were grouped into the following categories: neural networks; case-based reasoning; rough set based techniques; operational research; evolutionary approaches; decision trees; fuzzy logic, support vector machine and isotonic separation; and soft computing. All different groups were reviewed in detail with significant emphasis on the source of data, financial ratios applied, country of origin, time period, and relative performance in terms of prediction accuracy for the sample studied (Kumar & Ravi, 2007).

Another thorough review of intelligent techniques used in failure prediction was presented by Sun, Li, Huang, and He (2014) focusing on the definition of financial distress, modelling, sampling approaches and featuring approach. The review paper discussed the following intelligent techniques: artificial intelligence; hybrid methods based on two or three algorithms; ensemble methods; dynamic models; and decision implementation methods.

Intelligent techniques were compared to statistical techniques by Le and Viviani (2017) in determining failure prediction of US banks using MDA and LRA models. These were compared to three intelligent techniques: artificial neural networks (ANN); k-nearest neighbour; and support vector machines. To determine the best financial ratios for failure prediction, 31 ratios were calculated from financial data up to 5 years prior to failure. The five techniques were applied to a sample of 1438 failed and 1562 non-failed US banks collected between 2008 and 2014. The results of the study showed that artificial neural networks and k-nearest neighbour techniques perform more effectively than the statistical models. However, logistic regression prediction accuracy was comparable with the artificial neural network and k-nearest neighbour methods. The support vector machine technique was outperformed by both statistical models. Artificial neural networks and k-nearest neighbour proved effective in correctly detecting failure when the other methods could not. Based on the results, the authors concluded that LRA analysis can be used effectively to determine failure and complemented by artificial neural networks and k-nearest neighbour to classify the most difficult cases.

Considering the numerous financial ratios that can be calculated from financial data, most studies select variables based on statistical considerations. No theory has been proposed to indicate which financial ratios are the best variables to use as predictors. As such, researchers typically select variables based on statistical significance of the variable, the sign of the variables' coefficients, individual discriminating ability of each variable using a univariate analysis, factor analysis, and stepwise methods such as backward elimination and forward selection (Balcaen & Ooghe, 2006). To eliminate such methods of determining model independent variables, intelligent techniques have been implemented to determine such variables guided only by the sample data considered.

Acosta-González and Fernández-Rodríguez (2014) developed a Logit model by applying a genetic algorithm to choose a parsimonious model with the least number of financial ratios as predictors. The genetic algorithm was guided by a Schwarz information criterion to explore multiple models made available by all the possible existing financial ratios. A genetic algorithm in such an application was effectively an optimization technique based on principles of natural evolution. The empirical study focused on publicly listed Spanish firms in the building industry. In-sample and out-of-sample data was considered. Failure prediction models were estimated using an in-sample data consisted of 93 failed and 254 non-failed firms in 2004. These models were evaluated for predictive accuracy using an out-of-sample data set consisting of 80 failed and 320 non-failed firms between 2006 and 2009. For one-year prior to failure, the authors proposed the following Logit function:

$$L = -2.5658 - 0.0065\left(\frac{Profit}{Total\ assets}\right) - 0.0020(Working\ Capital)$$

Different variables were selected by the genetic algorithm for two-years, three-years, and four-years prior to failure. Out-of-sample overall classification accuracy was 83% for one-year prior to failure. The developed model was compared to Altman's z'-score model. For one-year ahead prediction, the z'-score model performed better at predicting failed firms, however, the authors' model better predicted non-failed firms. The authors' model clearly outperformed the Altman z'-model for two-, three-, and four-years ahead prediction.

Tian and Yu (2017) studied bankruptcy prediction in the international market with an empirical study of the Japan market. They applied the artificial intelligent technique of adaptive least absolute shrinkage and selection operator (LASSO) to determine a set of parsimonious financial ratios that can be used as predictors of bankruptcy. A hazard model, with the adaptively selected financial ratios, was applied to determine performance for time varying international market panel data over different prediction horizons. Empirical results for the Japan market show that bankruptcy can be accurately predicted over one-year, two-year, and three-year time horizons using three adaptively selected financial ratios. The statistically significant financial ratios for the Japan market, as selected by adaptive LASSO were: retained earnings to total assets; total debt to total assets; and current liabilities to revenue. The hazard model with these three variables correctly predicted 80% of bankruptcies for the one-year prior to failure horizon. When compared to the Altman (1968) z-score model, the adaptive LASSO hazard model achieved greater prediction accuracies. Prediction results for European countries such as UK, Germany and France were mixed.

## 2.4. Macroeconomic factors in failure prediction models

Rose, Andrews, and Giroux (1982) studied the effects of macroeconomic conditions on corporate failure. Their results showed that the significant macroeconomic variables were the: S&P Index; base interest rate; 3-month US Treasury Bonds rate; ratio of gross domestic private investment to GDP; and ratio of retail sales to GDP (Acosta-González, Fernández-Rodríguez, & Ganga, 2017). Lagged macroeconomic variables found to be significant predictors.

Altman (1983) examined general trends and macroeconomic conditions that affect corporate failures in the US. The study found that the probability of corporate failure is higher when there is a cumulative reduction in: real economic growth; stock market performance; money supply growth; and increased business formation.

Hol (2007) evaluated the failure of Norwegian firms based on financial statement analysis and an analysis of the macroeconomic environment. The GDP gap, an industrial production index and the money supply M1 were found to be significant predictors of bankruptcy probability. Their results showed that growth in economic activity was insignificant to bankruptcy probability.

Carling, Jacobson, Lindé, and Roszbach (2007) studied the impact of macroeconomic conditions on credit and default risk in Swedish firms. Three macroeconomic variables were considered in the study: output gap; yield curve; and house hold expectation. The output gap was included to indicate aggregate demand conditions, which if higher are expected to reduce default risk. The yield curve indicates future real activity and was estimated by taking the difference between annualized nominal interest rates on 10-year government bonds and 3-month treasury bill rates. Carling et al. (2007) applied an accounting based hazard model, with the three macroeconomic variables, to predict credit and default risk in 54 603 Swedish firms between 1994 and 2000. Results indicated that macroeconomic variables have significant explanatory power when used with accounting financial ratios.

Das, Hanouna, and Sarin (2009) examined firm distress using credit default swap spreads data as a measure of default risk. Credit default swap spreads are derivatives that offer protection from the event that a given firm defaults on its debt obligations. The spreads indicate the compensation that market participants require for holding firm equity at a risk. Das et al. (2009) studied three models: purely accounting based; purely market based; and combination of accounting and market data. All three models incorporated interest rate, market returns, and industry specific returns as macroeconomic variables. The interest rate used was the risk-free rate estimated from 3-month treasury bills and market returns were calculated from prior year return on the S&P 500. Using a sample of 2860 credit default swap spreads from 2001 to 2005, results showed that the accounting and market based models correctly predict 65% and 64%, respectively. Their comprehensive model consisting of both accounting and market data correctly predicted 72% defaults.

Christidis and Gregory (2010) developed a model for predicting failure of UK firms using a hazard model that contains both accounting and market-based data. The study incorporated macroeconomic variables to measure their impact on the model's predictive accuracy. The first macroeconomic variable accounted for the change in the return of the FTSE All Share Index over the previous twelve months. Interest rates in the economy was determined from the nominal three-month Treasury Bill Rate. The inflation rate in the economy was determined from the monthly change in the Retail Price Index. To indicate recession, an inverted yield curve was incorporated by including the term structure premium. For direct measures of economic activity, the 12-month change in the Long Leading Indicator and monthly change in the Industrial Production Index were included in the study. To develop the model, a sample consisting of 589 failed publicly listed UK firms was collected between 1978 and 2006. The results showed that the risk-free rate of interest, the term structure of interest rates, and inflation rate were significant variables that improved the prediction accuracy of their proposed model.

Tinoco and Wilson (2013) estimated a failure prediction model based on accounting data, market data, and macroeconomic variables. The study used a sample of 379 failed and 2641 non-failed publicly listed firms between 1980 and 2011. Data was collected for accounting, market, and macroeconomic variables which was used to calculate 130 variables. These were reduced, using univariate analysis and a Logit procedure, to four accounting ratios, two macroeconomic variables and four market variables. The two macroeconomic variables found to improve prediction accuracy were: 3-month inflation adjusted Treasury bill rate; and the retail price index as a measure of inflation. Hazard failure prediction models based on the ten accounting, market, and macroeconomic variables were analysed. The authors concluded that models based solely on financial data cannot definitively predict corporate failure. Thus, additional information such as macroeconomic variables can improve failure prediction of accounting-based models.

Alifiah (2014) applied a LRA model using financial ratios and macroeconomic variables to predict financial distress in Malaysian companies. The empirical study focused on companies in the trading and services sector. The macroeconomic environment effects of the base lending rate, consumer price index, gross domestic product, the Malaysian top 30 index (KLCI), and money supply (M2) were included as independent variables. The data used in the study consisted of 10 failed and 10 non-failed firms for the period from 2001 to 2010. From the five macroeconomic variables, the base lending rate was found to be statistically significant in predicting failure. The final model had a Logit with the following independent variables: debt ratio; total assets turnover; working capital ratio; net income to total assets ratio; and base lending rate. This model correctly classified 85% of the corporate failures in the sample.

Asgarnezhad Nouri and Soltani (2016) developed failure prediction models for non-financial firms in Cyprus. LRA was applied to accounting, market and macroeconomic variables to develop a prediction model. The macroeconomic variables, as recommended by Tinoco and Wilson (2013), of inflation rate, interest rate, and economic growth rate were studied. The inflation rate was determined from the Consumer Price Index. Interest rate was calculated from the 3-month interest rate on deposit from households. Economic growth rate was calculated as the annual change in gross domestic product. Accounting and market-based models correctly predicted failure 91% and 82% respectively. Macroeconomic variables were shown to be insignificant in predicting a firm's probability of failure.

In a recent study, Acosta-González et al. (2017) observed an increase in the rate of corporate failures in a time of economic crisis. The authors developed a Logit model in which both financial ratios and macroeconomic variables are used to predict corporate failure. Similar to the work of Acosta-González and Fernández-Rodríguez (2014), the variables in the Logit model were selected from a large set of financial ratios and macroeconomic variables using a genetic algorithm guided only by data. The study focused on the corporate failures in the Spanish construction industry between 1995 and 2011. For one-year prior to failure, the genetic algorithm selected the following macroeconomic variables as predictors: spread of the interest rates term structure; credit granted by banks to householders; price per urbanized square metre. The empirical results showed that the model successfully classified failed and non-failed firms with an accuracy of 98.5% and 82.5%, respectively, for one-year in prior to failure. For two-, three-, four-, and five-year forecast, other macroeconomic variables such as 12-month interest rate, stock market volatility, sector's share of gross domestic product were selected by the genetic algorithm.

## 2.5. Empirical Research

In the review papers of Kumar and Ravi (2007) and Adnan Aziz and Dar (2006), empirical studies in literature which examined statistical and intelligent techniques application to failure prediction to firms and banks in different countries were thoroughly detailed. Kumar and Ravi (2007) found that much of studies were conducted using data between 1980 and 2003 with a clear majority of based on USA data followed by European countries. The empirical studies presented in both review papers are not discussed in this report. Empirical studies proceeding both review papers are discussed in this section.

Reisz and Perlich (2007) found that for one-year prior to failure, Altman's (1968) z- and z"-score models outperformed their contingent claim model and their computationally intensive down-and-out barrier option model. However, for three-year and five-year forecasting the Altman z- and z"-score models performed poorly compared to the down-and-out barrier option model.

Agarwal and Taffler (2008) evaluated the performance of the accounting-based z-score model of Taffler (1984), specifically formulated for UK companies, against market-based option-pricing models for UK listed companies between 1985 and 2001. The z-score model performed marginally better than option-pricing models, however, the difference was statistically not significant. For credit risk assessment, the results revealed that banks could realise significantly higher risk-adjusted revenues, profits, return on capital, and return on risk adjusted capital if they applied the z-score model. Their results indicated that failure prediction models using accounting ratios are comparable to market-based option-pricing models.

Das et al. (2009) studied the difference between accounting-based and market-based models in pricing firm distress. A new accounting-based model was constructed using 10 variables which capture firm size, profitability, liquidity, trading account activity, sales growth, and capital structure. The market-based model studied was based on the Merton (1974) methodology for calculating probability of default. The two models were compared for pricing firm distress using credit default swaps which are derivatives that offer protection from the event of a firm defaulting on its debt obligations. The study concluded that the accounting-based model performed the same as market-based models. For some cases, the accounting-based model results were better.

Lin (2009) examined four models for predicting failure of Taiwan public industrial firms between 1998 and 2005. The four models studied were: MDA; LRA; Probit; and ANN. The models were validated using with sample tests and out-of-sample tests. The results indicated that the Probit, LRA, and ANN models achieved higher prediction accuracy. The Probit model was found to be the best and offered stable performance.

Wu, Gaunt, and Gray (2010) tested five bankruptcy models that are prominent in literature using a data set from 1980 to 2006 which included 887 bankruptcies. The models were as follows: z-score model of Altman (1968); Logit model of Ohlson (1980); Probit model of Zmijewski (1984); Hazard model of Shumway (2001); and BSM model of Hillegeist et al. (2004). The analysis found that the MDA model performed poorly compared to the other models. The Logit and Probit model accuracies deteriorated over time within the period investigated. The hazard model outperformed all models.

In their comparative study, Jackson and Wood (2013) compare accounting-based models with market-based contingent claims models based on data UK publicly listed firms. The accounting-based models were based on: univariate analysis; MDA; LRA; and artificial neural networks. Three univariate models were included in the analysis: cash flow to total debt; firm size; and book-to-market value ratio. Two forms of the MDA z-score models were analysed: Altman (1968) and Taffler (1983). The Logit model of Ohlson (1980) was tested. Neural networks methods were applied to create models based on the variables from Altman (1968), Ohlson (1980), and Taffler (1983). In total, 25 different methods were analysed and compared using ROC curves. The results of the analysis show that accounting-based models were outperformed by market-based contingent claims models. The univariate model of firm size performed better than MDA and LRA models. The univariate model of cash flow to total debt, which was identified by Beaver (1966), was found to have similar predictive accuracy to MDA models. The inferior performance of accounting-based models was attributed to the sample period, from 2000 to 2009, which included the global financial crisis.

A comparative study conducted by Bauer and Agarwal (2014) on UK publicly listed firms compared hazard models, using both accounting and market data, to accounting-based and contingent claims models. The hazard model of Shumway (2001) performed better than the Taffler (1983) z-score and the naïve contingent claims model of Bharath and Shumway (2008). Economic value of the models was evaluated considering that lending to a firm that fails results in a higher cost than the opportunity cost of not lending to a firm that does not fail. Such an assessment was argued to be true for competitive loan markets with different costs of misclassification. The hazard model of Shumway (2001) resulted in the highest economic benefit.

Castagnolo and Ferro (2014) examined models based on MDA, LRA, BSM based contingent claims models. The results showed that the Logit model of Ohlson (1980) outperformed the other models studied. The BSM based contingent claims model was found to be insufficient as a default predictor. The authors propose that a more accurate default prediction model can be built by combining information from different prediction models containing both accounting and market data.

To develop a bankruptcy prediction model for the Slovak Republic, Mihalovic (2016) considered and compared two models. The two models were estimated using MDA and LRA. To estimate both models, a sample of 118 failed and 118 non-failed firms was collected for the year 2004. The results showed that the Logit model outperformed the classification accuracy of the MDA model.

Most failure prediction models were estimated and performed well for developed markets such as US and UK firms. Charalambakis and Garrett (2016) investigated accounting-based and market-based failure prediction models for the UK, a developed market, and India, an emerging market. For the UK, the study found that a hazard model which combines accounting and market data outperformed the Taffler (1983) z-score model. However, this model performed poorly when applied to Indian firms where market data fails to predict bankruptcy. In India, an accounting-based model performed best in predicting corporate failure. In conclusion, the study found that accurate failure prediction models from developed economies cannot be successfully applied to emerging economies.

## 2.6. South Africa failure prediction studies

Corporate failure prediction literature in South Africa was comprehensively reviewed by Naidoo and Du Toit (2007). The review reported on research conducted from 1977 to 1999 which included five models based on univariate analysis, MDA, and Bayes-Fisher discrimination. These five studies are not included in this literature review. Steyn-Bruwer and Hamman (2006) presented a comparison of 8 South African studies conducted between 1990 and 1995. Research on failure prediction in South Africa since 2004 are detailed in this section to present the state of literature in the country.

Kidane (2004) performed a thorough investigation into the applicability of the Altman (2000) and Springate (1978) z-score models in predicting financial failure of South African companies. The study aimed to test both models for practical applicability and relevance in South Africa as the models were developed in different economic environments and using data from other countries. The applicability of both models was empirically tested using information technology and services companies listed on the Johannesburg Stock Exchange (JSE). The data sample consisted of 24 failed and 62 non-failed that were listed on the JSE from 1999 to 2003 with a mean turnover of R1 051 million, ranging from R0.150 million to R20 677 million. For the failed companies, annual financial statements up to five years prior to failure were used. The z-score models studied failed to correctly predict the failed and non-failed companies in the data sample. Therefore, the study concluded that both models should not be used to predict bankruptcy in the information technology and services sector.

Steyn-Bruwer and Hamman (2006) used the intelligent technique of recursive partitioning, which is a classification tree algorithm, to develop a failure prediction model for the South African environment. The study defines corporate failure as a condition when a company cannot survive in its existing structure which encompasses delisting or major structural change. A company was identified as failed if it failed within four years after year-end, otherwise it was classified as non-failed. A data sample of JSE listed industrial companies between June 1997 and May 2002 was used in developing the model. The resulting model was tested on an independent sample to report predictive accuracy and not the classification accuracy. Classification accuracy, which indicates the model's accuracy in predicting the development population data, was identified as inadequate as it will always be high. The research identified three important variables as predictors of firm failure: the size of a firm measured by total assets; cash flow from operating activities divided by revenue from the previous financial year; and cumulative three-year period cash flow from activities divided by cumulative three-year period revenue.

Naidoo and Du Toit (2007) proposed a two-stage approach to determine the financial health of South African firms. In the first stage, models were developed to predict the financial health of a firm. In the second stage, the authors developed a methodology to provide information, independent of the first stage, to present a meaningful picture of the company to stakeholders. To better represent the continuum of financial failure, the authors do not use the traditional dichotomous definition of failed and non-failed, but rather present three states of health as follows:

- Healthy if real earnings growth is positive
- Intermittent if real earnings growth is negative
- Distressed if earnings are negative

For the first stage, Naidoo and Du Toit (2007) used a sample consisting of 42 firms divided into a test and holdout sample. The test sample consisted of firm data from 1970 to 1979, whereas the holdout sample covered the period from 1970 to 1999. To develop three statistical models, a total of 34 variables was considered: 26 financial ratios; 6 macroeconomic variables; and 2 lagged macroeconomic variables. Results showed that the two univariate models proposed performed better than the 2-variable MDA model they developed.

Muller, Steyn-Bruwer, and Hamman (2009) extended the work of Steyn-Bruwer and Hamman (2006) by studying other modelling techniques to compare prediction accuracies using the same sample data set. The failure prediction modelling techniques considered were: MDA; recursive partitioning; LRA; and Neural Networks (NN). The study defined the cost of failure related to Type I and Type II errors using a what they termed the Normalised Cost of Failure (NCF). The NCF considers that the typical cost of Type I errors is between 20 to 38 times that of Type II errors. Results showed that different modelling techniques produced different predictive accuracies. The MDA and recursive partitioning models correctly predicted the most number of failed companies which resulted in lower NCF values. The LRA and Neural Networks models provided the best overall predictive accuracy but with higher NCF values.

Hlahla (2011) studied failure prediction of JSE listed industrial firms excluding the financial services sector. The sample date consisted of 14 failed and 14 non-failed companies compiled from Sharenet and Stock Exchange News Service (SENS). Balance sheet, income statements, and cash flow statement data was collected for all firms in the sample five years prior to financial distress. Financial statements were collected in a standardised format from the BFA McGregor information portal. The non-failed firms in the data were matched by fiscal years, industry, and asset size to those of the failed firms. Matching of failed and non-failed firms in the sample data minimises bias in the prediction model and is consistent with the previous studies of Beaver (1966) and Altman (1968). To avoid distortion of the data sample, the data were collected between 2003 and 2007, excluding failed firms from 2008 onwards as these could have been negatively affected by the global financial crisis. The average total asset size of failed companies was R165 million and for non-failed companies was R252 million. MDA was chosen as the prediction model to create a z-score function. The initial list of independent variables consisted of 64 financial ratios selected based on importance in assessing failure or success of a company. Cash flow ratios as well as ratios with the potential to be negative were excluded. The independent variables were tested for their discriminating power using a stepwise discriminant model and normality test. According to Altman (1968), the MDA requires independent variables that are normal. None of the independent variables were found to be normal, 26 variables were log normal and one was square root normal. Independent variables found to not be normal were excluded from the study. Thus, 26 financial ratios were selected as independent variables for the Multiple Discriminant Analysis. The result of the MDA indicated that the most significant discriminant independent variables were: times interest earned ratio; cash-to-debt ratio; and working capital turnover. Hlahla determined that firms with a z-score below -0.493 were distressed, while firms with a z-score above 0.525 were non-distressed. The grey area between z-score -0.493 and 0.525 was classified into two categories. The first category, z-score from -0.493 to 0.016 is classified as the grey area of distressed firms. The second category, z-score from 0.016 is classified as the grey are of non-distressed firms. For one year prior to failure, the developed model predicted failure with 100% accuracy for the sample data. However, the accuracy rate deteriorated from two years up to four years prior to failure. Thus, the overall classification accuracy of the model was calculated as 75.3%. The model was tested out-of-sample with of 8 failed companies not included in the original sample and correctly classified all failures in the year of distress.

Masekesa (2011) examined corporate failure rate in small medium and micro enterprise companies in South Africa to determine macroeconomic risk factors related to corporate failure. Their study focused on corporate failures between 1994 and 2009 which included the global financial crisis of 2007-2009. Results indicated that corporate failure rates in South Africa are significantly and positively related to the average lending rate, inflation rate, new corporation, exchange rate, the global financial crisis. Corporate failure rate was shown to be inversely related to gross domestic product and money supply. The author concluded that a consistent relationship between the macroeconomic environment and corporate failures exists.

Rama (2012) studied the applicability of Altman's (1968) z-score model to South African JSE listed companies. The model was tested on 227 companies using data from the 2008 financial year, however, companies in the financial service, mining sectors and the top 40 JSE listed companies were excluded from the sample. The author argued that the financial and mining sectors a highly specialised and highly regulated to the extent that they differ from other sectors when comparing profitability and asset structures. The top 40 JSE listed companies were excluded based on their size and history lowering the probability of such companies experiencing financial distress. The Altman z-score was calculated using 2008 data and success or failure of the firms was measured in 2009 and 2010. The results show that the z-score model is applicable to South African companies when z-scores are positive and do not fall into the range of uncertainty or "grey area" specified by Altman (1968). When z-scores are negative, the model was found to be unreliable.

Cassim (2014) investigated the use of univariate analysis in predicting failure of JSE listed companies. The author applied two univariate analyses techniques: comparative financial ratio analysis; and ratio trend analysis. The univariate models were compared to Altman's z"-score model. The study analysed 8 failed and 8 non-failed companies with financial statement data collected from 2007 to 2012. Financial and mining sectors we excluded from the data set based on the same reasoning as Rama (2012). Results indicate that the Altman z"-score model classifies companies as follows:

- z"-score > 2.60 is financially healthy and stable (non-failed)

- z”-score between 1.10 and 2.60 is the area of uncertainty
- z”-score < 1.10 is failed

The author concluded that a combination of various prediction models could enhance failure prediction accuracy.

A summary of the South African literature on failure prediction is presented in Table 2, with information on the period, industries, sample data, and source of data for each study.

*Table 2: South African literature data*

| Study | Period | Industries Covered | Sample Data | Data Source |
|---|---|---|---|---|
| Kidane (2004) | 1999 – 2003 | Services, Information Technology | 86 JSE listed (24 failed) (62 non-failed) | Bureau of Financial Analysis SA |
| Steyn-Bruwer and Hamman (2006) | 1995 – 2002 | Industrial | JSE listed | Not specified |
| Naidoo and Du Toit (2007) | 1970 – 1999 | Various | 42 JSE listed (20 failed) (22 non-failed) | McGregor BFA |
| Hlahla (2011) | 2003 – 2007 | Not specified | 28 JSE listed (14 failed) (14 non-failed) | Sharenet & SENS |
| Rama (2012) | 2008 | All industries, excluding mining and finance | 227 JSE listed | McGregor BFA |
| Cassim (2014) | 2007 – 2012 | Consumer Goods, Consumer Services, Industrials | 16 JSE listed (8 failed) (8 non-failed) | McGregor BFA |

## 2.7. Summary of Literature Review

Since the first prediction model by Beaver (1966), many researchers have studied corporate failure and proposed numerous models. Most of the work in literature is focused on developed economies such as US and UK. Results in literature show that such models cannot be simply applied to developing economies as the predictors of failure or predictor coefficients may differ. Various researchers in South Africa have studied failure prediction and proposed univariate or MDA models. Masekesa (2011) studied the macroeconomic factors that can increase or decrease the corporate failure rate to guide policymakers. This study is undertaken to fill a gap in South African literature by using a logistic regression model which incorporates financial ratios, firm characteristics, macroeconomic variables, and industry sector characteristics in predicting failure.

# Chapter 3 - Research Methodology

## 3.1. Introduction

This section describes the research methodology applied in this study. The research hypotheses which need to be tested are clearly defined. The research design, data, data source, methodology and the econometric model to be applied are explained to formulate how the hypotheses will be tested.

## 3.2. The hypotheses

The research hypotheses are developed based on the literature survey and the aim of this study. The benchmark for this study is the research of Altman et al. (2017) that developed multiple logit models based on international data that excluded South Africa. Four hypotheses are proposed to test the accounting variable only model of Altman et al. (2017) with South African firm data and extend the work by including macroeconomic environment effects.

### 3.2.1. H1: Re-estimation of model coefficients

The logit model of Altman et al. (2017) based on the z”-score model was developed using international data from mostly developed economies. As discussed in the literature review, Agarwal and Taffler (2007) cautioned against using models developed using data from other countries while Charalambakis and Garrett (2016) found that models from developed economies cannot be applied to emerging economies. Grice and Dugan (2003) showed that re-estimating model coefficients with country specific and up to date information improved model accuracy. Based on these previous studies, it is hypothesised that re-estimating the coefficients of the Altman et al. (2017) logit model based on four accounting variables, using South African firm data, will improve the classification performance of the model. The first hypothesis is as follows:

> **H1**: Re-estimating the coefficients of the Altman et al. (2017) logit model based on the four z”-score accounting variables improves its prediction accuracy

### 3.2.2. H2: Firm size

Larger firms are believed to have a lower probability of experiencing financial distress and eventual failure (Rama, 2012). Altman et al. (2017) found that the size effect has a significant impact on failure prediction model accuracy. Thus, firm size needs to be studied in the context of failure prediction models. The second hypothesis assumes that the classification performance of the benchmark model, which is based on the four z”-score accounting variables, is improved when the size of the firm is explicitly incorporated. The second hypothesis is as follows:

> **H2:** Model prediction accuracy is higher when the effect of firm size is included

### 3.2.3. H3: Firm age

The risk of failure is known to be a function of the age of a firm. Older firms have a lower risk of failure, while very young firms typically have a very high risk of failure. Young firms are more likely to suffer from capability and resource deficits than older firms which have developed valuable capacity for both (Pretorius, 2008). Altman et al. (2017) found that young firms have a higher risk of failure and that including the effect of age marginally improved model performance. Thus, the age of a firm needs to be studied in the context of South African firm failure prediction. The third hypothesis assumes that the classification performance of the benchmark model will improve when the age of the firm is explicitly incorporated into the model. The third hypothesis is as follows:

> **H3:** Model prediction accuracy is higher when the effect of firm age is included

### 3.2.4. H4: Macroeconomic environment

The literature supports the notion that corporate failure is influenced by the macroeconomic environment in which companies operate. The macroeconomic variables of economic growth (Altman, 1983), inflation (Tinoco & Wilson, 2013), and lending rate (Alifiah, 2014) have been studied extensively and found to be significant predictors by other international literature. In South Africa, Masekesa (2011) found all three factors to be significant predictors of corporate failure. The forth hypothesis assumes that considering macroeconomic variables specific to South Africa will improve the classification performance of the benchmark model, which is based on the four z”-score accounting variables. The forth hypothesis is as follows:

> **H4:** Model prediction accuracy increases when the effect of the macroeconomic environment is included

### 3.2.5. H5: Sector of the firm

Firms in different sectors, by nature of their business and how revenues are generated, could present different values for the same financial ratios. A study by Smith and Liou (2007), which refers other previous studies, found that firm failure is influenced by the sector effect. Altman et al. (2017) investigated the effect of different industries with results showing that firms in the construction industry are at a higher risk of failure, followed by the manufacturing industry. The fifth hypothesis assumes that considering the South African industries in which firms operate will improve the classification performance of the benchmark model, which is based on four accounting variables. The fifth hypothesis is as follows:

> **H5:** Model prediction accuracy is higher when the industry or sector in which firms operate is included

## 3.3. Research design

The aim of this study is to assess financial statements of firms and the macroeconomic environment in which they operate to predict corporate failure in a developing economy such as South Africa. The study focuses on healthy and failed companies listed on the JSE between 2000 and 2015. Macroeconomic variables, firm characteristics and financial ratios calculated from financial statement data are used to investigate the corporate failure phenomenon. A cross-sectional logit model is estimated by considering financial, structural and macroeconomic explanatory variables to predict corporate failure. Thus, the considered logit is of the form:

$$L = b_0 + \sum_{i=1}^{4} b_i x_i + b_5 \ln(TA) + b_6 AGE + b_7 \Delta GDP + b_8 \Delta CPI + b_9 PLR + \sum_{j=1}^{3} c_j D_j$$

where $L$ is the logit; $x_1$ is the ratio of working capital to total assets (WCTA); $x_2$ is the ratio of retained earnings to total assets (RETA); $x_3$ is the ratio of earnings before interest and tax to total assets (EBITTA); $x_4$ is the ratio of the book value of equity to total debt (BVETD); firm size is measured as the natural logarithm of total assets (LNTA); $AGE$ is the age of the firm in years since inception; $\Delta GDP$ is the annual gross domestic product (GDP) growth rate; $\Delta CPI$ is the annual inflation rate as measured by the change in consumer price index (CPI); $PLR$ is the prime lending rate; and $D_j$ are dummy variables for different industry sectors.

The constant $b_0$ is the intercept of the logit and $b_i$ $(i = 1, \dots, 9)$ are coefficients of the independent variables and $c_j$ $(j = 1, \dots, 3)$ are coefficients of sector dummy variable.

## 3.4. Data and data source

Firms that conform to this study's definition of failure were identified using JSE information portals, Sharedata and SENS. A comprehensive list of companies that were delisted from the JSE between 2000 and 2015 was compiled. Those companies that delisted for reasons such as voluntary winding up, non-compliance with JSE listing requirements, conversion, part of a scheme arrangement or mergers were eliminated. The list was reduced to 53 companies. More companies were eliminated from the list according to the delimitations set out in chapter 1. The final list consisted of 34 failed firms which conform to this study's definition of failure. Financial statements of these firms were collected for ten-years prior to failure from the INET BFA Expert (formerly known as McGregor BFA) platform. The INET BFA Expert platform is a financial reporting database which provides historical and fundamental data on South African and international listed and delisted companies. The financial statements were collected in a standardized format which included balance sheet, income statement, and cash flow statement data for each firm in the sample.  Failed companies identified from the JSE information portals are presented in Table 3 with their corresponding year of delisted from the JSE.

*Table 3: Sample of failed firms and year of delisting from the JSE*

| **COMPANY (CODE)** | **YEAR** | **COMPANY (CODE)** | **YEAR** |
|---|---|---|---|
| 1TIME HOLDINGS LTD (1TM) | 2011 | GLOBAL TECHNOLOGY LTD (GLT) | 2003 |
| AFRICA CELLULAR TOWERS LTD (ATR) | 2011 | HICOR LTD (HOR) | 2001 |
| AG INDUSTRIES LIMITED (AGI) | 2009 | KAIROS INDUSTRIAL HOLDINGS LIMITED (KIR) | 2011 |
| ALERT STEEL HOLDINGS LTD (AET) | 2013 | MB TECHNOLOGIES LTD (MBT) | 2001 |
| AQUA ONLINE HOLDINGS LTD (AQU) | 2003 | NORTHERN ENGINEERING IND AFRICA LD (NEI) | 2000 |
| BEGET HOLDINGS LIMITED (BEE) | 2009 | PALS HOLDING LIMITED (PAL) | 2007 |
| BEST CUT LIMITED (BCH) | 2008 | PAMODZI GOLD LTD (PZG) | 2007 |
| BILLBOARD HOLDINGS LTD (BLL) | 2000 | PLATFIELDS LTD (PLL) | 2012 |
| BIOSCIENCE BRANDS LTD (BIO) | 2012 | PROTECH KHUTHELE HOLDINGS LTD (PKH) | 2013 |
| CCI HOLDINGS LIMITED (CCG) | 2002 | SANYATI HOLDINGS LTD (SAN) | 2011 |
| CHEMICAL SPECIALITIES LTD (CSP) | 2014 | SERVEST HOLDINGS LTD (SRV) | 2002 |
| CORE HOLDINGS LTD (COR) | 2001 | SILTEK LTD (LIQUIDATION) (STK) | 2000 |
| DIALOGUE GROUP HOLDINGS LIMITED (DLG) | 2010 | SOUTHERN MINING CORPORATION LTD (SMC) | 2003 |
| DUNLOP AFRICA LTD (DNL) | 2000 | SQUARE ONE SOLUTIONS GROUP LTD (SQE) | 2008 |
| ERBACON INVESTMENT HOLDINGS LTD (ERB) | 2013 | TOP INFO TECHNOLOGY HOLDINGS LTD (TOT) | 2000 |
| FARITEC HOLDINGS LTD (FRT) | 2009 | UNIVERSAL GROWTH HOLDINGS LTD (UNG) | 2000 |
| FASHION AFRICA LTD (FSH) | 2002 | WHETSTONE INDUSTRIAL HOLDINGS LTD (WTS) | 2000 |

The financial statements of non-failed firms were collected from 2000, or the year the firm listed on the JSE if it was after 2000, to 2015.  Non-failed companies were selected from all JSE sectors which were not excluded according to the delimitations stated in Chapter 1.  In total, 65 non-failed firms were identified.  Unlike the MDA method which requires the number of failed and non-failed firms be the same, LRA does not require matched samples.  The formulation and analysis of the logit model allows disproportional samples (Balcaen & Ooghe, 2006).  The year in which non-failed firms were analysed was selected randomly.  The final sample consisted of 99 firms, of which 34 had failed and 65 non-failed firms.

*Table 4: Summary of macroeconomic time series data*

| **Year** | **ΔGDP** (%) | **ΔCPI** (%) | **PLR** (%) |
|---|---|---|---|
| 2000 | 4.16 | 5.39 | 14.50 |
| 2001 | 2.74 | 5.64 | 13.00 |
| 2002 | 3.67 | 9.15 | 17.00 |
| 2003 | 2.95 | 5.87 | 11.50 |
| 2004 | 4.56 | 1.43 | 11.00 |
| 2005 | 5.28 | 3.35 | 10.50 |
| 2006 | 5.60 | 4.62 | 12.50 |
| 2007 | 5.36 | 7.15 | 14.50 |
| 2008 | 3.19 | 10.99 | 15.00 |
| 2009 | -1.54 | 7.12 | 10.50 |
| 2010 | 3.04 | 4.26 | 9.00 |
| 2011 | 3.28 | 4.99 | 9.00 |
| 2012 | 2.21 | 5.62 | 8.50 |
| 2013 | 2.49 | 5.76 | 8.50 |
| 2014 | 1.70 | 6.09 | 9.13 |
| 2015 | 1.30 | 4.58 | 9.42 |

Data for the economic growth rate, inflation and lending rate were collected from the International Monetary Fund (IMF) economic database. This data was collected as an annual time series over the period 2000 to 2015. Where data was collected monthly, the annual average was calculated. The macroeconomic time series data is shown in Table 4 with all variables are expressed in percentages.

To validate the performance models that will be developed on the sample from 2000 to 2015, it is necessary to collect out-of-sample data. Data for firms identified as failed based on financial statements in 2016 will be used for out-of-sample tests. No firms were found to have failed in 2016 as per definition of failure adopted in this study. In the absence of failed firms, data was collected for firms in financial decline. Firms were identified as financially decline if they had:

- Total liabilities exceeding total assets in 2016; or
- Two consecutive years of negative net income (2015 and 2016)

The second criteria of two consecutive years negative net income was applied by Tinoco and Wilson (2013). Using both definitions for distressed firms, 6 firms were identified and financial statements collected using the INET BFA Expert platform. The list of distressed firms is shown in Table 5 with corresponding cause of distress.

*Table 5: List of distressed firms in 2016*

| COMPANY (CODE) | Distress Condition |
|---|---|
| ARCELORMITTAL SA LTD (ACL) | Consecutive negative net income |
| BUFFALO COAL CORPORATION (BUC) | Total liabilities exceed total assets |
| IMBALIE BEAUTY LTD (ILE) | Consecutive negative net income |
| MINE RESTORATION INVESTMENTS LTD (MRI) | Total liabilities exceed total assets |
| PSV HOLDINGS LTD (PSV) | Consecutive negative net income |
| W G WEARNE LTD (WEA) | Consecutive negative net income |

Financial statement data was collected for 11 healthy firms for the 2016 financial year. Non-failed or healthy firms were identified as those firms whose total liabilities did not exceed their total assets. These were firms not included in the original sample data. In total, the out-of-sample data consisted of 6 distressed and 11 healthy firms presented in Table 6.

*Table 6: Out-of-sample list of firms*

| Company (Code) | Financial Condition |
|---|---|
| ACCENTUATE LTD (ACE) | Healthy |
| AECI LTD (AFE) | Healthy |
| CURRO HOLDINGS LTD (COH) | Healthy |
| DATACENTRIX HOLDINGS LTD (DCT) | Healthy |
| HOLDSPORT LTD (HSP) | Healthy |
| HUDACO INDUSTRIES LTD (HDC) | Healthy |
| QUANTUM FOODS HOLDINGS LTD (QFH) | Healthy |
| RHODES FOOD GROUP HOLDINGS LTD (RFG) | Healthy |
| THE SPAR GROUP LTD (SPP) | Healthy |
| TRUWORTHS INTERNATIONAL LTD (TRU) | Healthy |
| YORK TIMBER HOLDINGS LTD (YRK) | Healthy |
| ARCELORMITTAL SA LTD (ACL) | Distressed |
| BUFFALO COAL CORPORATION (BUC) | Distressed |
| IMBALIE BEAUTY LTD (ILE) | Distressed |
| MINE RESTORATION INVESTMENTS LTD (MRI) | Distressed |
| PSV HOLDINGS LTD (PSV) | Distressed |
| W G WEARNE LTD (WEA) | Distressed |

## 3.5. Econometric Logit model

### 3.5.1. Logistic Regression Analysis

In failure prediction, the dependant variable for analysis is the financial status of each individual firm, $Y_i$. In a dichotomous analysis, the problem is formulated such that $Y_i = 1$ for a failed firm and $Y_i = 0$ for a non-failed firm. The dependant variable $Y_i$ can be formalized and analysed based on its distribution conditional on the independent variables such that: the probability of failure ($Y_i = 1$) is $P_i$; and the probability of non-failure ($Y_i = 0$) is $1 - P_i$. That is,

$$Y_i = \begin{cases} 1 \\ 0 \end{cases} for \quad \begin{matrix} P_i \\ 1 - P_i \end{matrix}$$

where $P_i = P(Y_i = 1|x)$ is a logistic distribution function of the explanatory or independent variables $x_i$ (Gujarati & Porter, 2009). That is,

$$P(Y_i = 1|x) = \frac{1}{1 + e^{-L_i}} = \frac{1}{1 + e^{-(b_0 + \sum b_i x_i)}}$$

where $b_i$ $(i = 0,1, \dots, n)$ are the coefficients, $x_i$ $(i = 0,1, \dots, n)$ are the independent variables and $L$ is called the logit, hence the name logit model. Manipulation of the logistic distribution function defines the odds ratio:

$$\frac{P_i}{1 - P_i} = e^{L_i}$$

The odds ratio represents the ratio of the probability that a firm will fail to the probability that it will not. Taking the natural log of the odds ratio defines the logit as:

$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = b_0 + \sum b_i x_i$$

The logit model has the following features:

- when $P$ varies from 0 to 1, the logit $L$ varies from $-\infty$ to $+\infty$
- the logit $L$ is linear in the independent variables, $x$, while the probabilities are not
- if the logit $L$ is positive, an increase in the value of the independent variables increases the odds that $Y$ equals 1
- if the logit L is negative, an increase in the value of the independent variables decreases the odds that $Y$ equals 1
- assumes that the log of the odds ratio is linearly related to $x_i$

Typically, a probability $P_i \geq 0.5$ indicates a high probability a firm will fail, whereas $P_i < 0.5$ indicates a healthy firm.

To determine the coefficients $b_i$ of the logit, the method of maximum likelihood is used (Gujarati & Porter, 2009). In the method of maximum likelihood, the probability that $Y_i = 1 \; or \; 0$ is donated as $f_i(Y_i)$ such that the log likelihood function can be defined as:

$$\ln f_i(Y_1, Y_2, \dots, Y_n) = \sum_1^n Y_i \left(b_0 + \sum b_i x_i\right) - \sum_1^n \ln\left(1 + e^{(b_0 + \sum b_i x_i)}\right)$$

which is a function of the Logit coefficients $b_i$ $(i = 0,1, \dots, n)$. In the maximum likelihood method, the objective is to maximise this log likelihood function. To maximise the log likelihood function, the function is differentiated with respect to the logit coefficients and solved using methods of nonlinear estimation (Gujarati & Porter, 2009).

The method of maximum likelihood results in asymptotic estimated standard errors for which the statistical significance of coefficients can be evaluated using the standard normal $Z$ statistic. To test the null hypothesis that all the slope coefficients are equal to zero, the likelihood ratio statistic is compute. The likelihood ratio test is equivalent of the $F$ test in linear regression modelling. Given the null hypothesis, the likelihood ratio statistic follows the $\chi^2$ distribution with degrees of freedom equal to the number of independent variables, excluding the intercept term or constant $b_0$ (Gujarati & Porter, 2009).

### 3.5.2. Measure of model accuracy

The performance of a prediction model in discriminating or classifying failed and non-failed firms is presented in a classification matrix shown in Table 7.

*Table 7: Classification matrix*

| **Actual Group** | **Predicted Group** | | **Total** |
|---|---|---|---|
| | *0 (non-failed)* | *1 (failed)* | |
| *0 (non-failed)* | True Negative (*TN*) | False Positive (*FP*) | Number of non-failed (*NN*) |
| *1 (failed)* | False Negative (*FN*) | True Positive (*TP*) | Number of failed (*NF*) |
| | | **Total Sample** | Total number of firms (*N*) |

A simple method of measuring a model's accuracy is to determine the percentage of firms correctly classified as the ratio of correct classifications to the total sample, such that:

$$Accuracy = \frac{TN + TP}{N}$$

Such a measure of accuracy is accompanied with Type I and Type II errors. A Type I error occurs when a failed firm is classified as non-failed and a Type II error occurs when a non-failed firm is classified as failed.

$$Type\ I\ error = \frac{FN}{NF}$$

$$Type\ II\ error = \frac{FP}{NN}$$

Two different models can yield equal accuracies but perform differently with respect to the types of correct and incorrect predictions they provide. Incorrect predictions from one model might be almost all false negative or misses, while those from the other model might be almost all false negative or false alarms. Due to this phenomenon, the usefulness of two such models in predicting failure could be quite different under various conditions. To overcome such limitations, the receiver operating characteristic (ROC) curve is a simple tool used to indicate all possible combinations of the relative frequencies of the various kinds of correct and incorrect predictions. The ROC methodology considers two kinds of accuracies: sensitivity and specificity. Sensitivity represents actual positive cases and specificity represents actual negative cases (Metz, 1978).

In ROC analysis, the False Positive Rate (FPR) and True Positive Rate (TPR) are calculated based on the classification matrix in Table 7. The FPR and TPR are defined as:

$$FPR = \frac{number\ of\ false\ positve\ predictions}{number\ of\ actual\ nonfailed\ firms} = \frac{FP}{NN}$$

$$TPR = \frac{number\ of\ true\ positve\ predictions}{number\ of\ actual\ failed\ firms} = \frac{TP}{NF}$$

The values of FPR and TPR are calculated for a single decision threshold that separates the probability of failure and non-failure, typically chosen to be 0.5. These values would constitute a single point on the ROC curve which has FPR on the $x$-axis and TPR on the $y$-axis. When the decision threshold is varied from 0 to 1 several times, corresponding multiple points of (FPR, TPR) are obtained and plotted to create an ROC curve (Bradley, 1997). According to Metz (1978), the ROC curve describes the inherent predictive characteristics of a model and the receiver of the model's information can operate at any point along the curve by using an appropriate decision threshold. The axes of the ROC graph both range from 0 to 1 as these are the limits of FPR and TPR values.

From the ROC curve, the measure of classification performance or accuracy computed is the area under the ROC curve (AUC). The AUC measures discrimination, which is the ability of a model to correctly classify failure and non-failure of firms (Kleinbaum & Klein, 2010). A ROC curve that is an upward sloping line from the bottom left corner (FPR = 0, TPR = 0) to the top right corner (1, 1), the AUC is 0.5 which indicates a 50% chance that a firm is failed or non-failed. Thus, the AUC increases as the curve approaches the top left corner where FPR is 0 and TPR is 1. A model with an AUC of 0.5 is like using a coin toss to predict failure or non-failure of a firm. The accuracy of a models with AUCs between 0.5 and 0.7 is low; between 0.7 and 0.9 the accuracy is moderate; and for over 0.9 the accuracy is high (Streiner & Cairney, 2007).

### 3.5.3. Multicollinearity Check

Multicollinearity in samples occurs when there is a perfect or linear relationship between two or more explanatory variables of a regression model. This arises when some of the independent variables measure the same concept or are linear combinations of other variables. Multicollinearity results in unstable model estimations and inflated standard errors (Tinoco & Wilson, 2013). When multicollinearity is present, the instability of coefficients estimates increases. No unique method of detecting it or measuring its strength have been developed, only rules of thumb. This study will use the variance inflation factor (VIF) as an indicator of multicollinearity. VIF shows how the variance of an independent variable is inflated by the presence of multicollinearity. As a rule of thumb, a variable with a VIF exceeding 10 is highly collinear (Gujarati & Porter, 2009).

### 3.5.4. Marginal effects

In OLS regression models, the slope coefficient measures the change in the average value of the dependent variable for a unit change in the value of an independent variable, with all other variables held constant (Gujarati & Porter, 2009). Coefficients of conditional probability models such as the logit, cannot be interpreted the same as OLS regression model slope coefficients. The effect of each variable on the probability of failure is interpreted by computing marginal effect. The marginal effect of an independent variable is defined as the partial derivative of the failure probability, $P(Y_i = 1|x)$, with respect to the independent variable of interest (Tinoco & Wilson, 2013). Recalling that the conditional probability of failure or logistic distribution function for a logit model is defined as:

$$P(Y_i = 1|x) = \frac{1}{1 + e^{-L_i}} = \frac{1}{1 + e^{-(b_0 + \sum b_i x_i)}}$$

The marginal effect of an independent variable is mathematically expressed as $\partial P / \partial x_i$.

## 3.6. Methodology

The five research hypotheses in this study will be tested by computation of model accuracy in discriminating failed firms from non-failed firms. The statistical analysis, computed using Gretl software, will be based on models consisting of financial ratios and macroeconomic data for each firm. The models will be tested for their predictive performance in determining corporate failure for one-year prior. For this, it is critical to distinguish between the year of failure and the last year in which financial statements were issued.

In his seminal work, Ohlson (1980) found the average lead time between the last financial statement date and bankruptcy was approximately 13 months. Agarwal and Taffler (2007) found that for US firms, the average lead time to failure from the date of the last financial statements was 13.2 months. For UK firms, Tinoco and Wilson (2013) found the average lead time between date of failure and the last financial statements was 14.2 months. Altman et al. (2017) state that the average lead time to failure is approximately 1 to 2 years from the date of the last financial statements. This study adopts the convention that the last financial statements for failed firms occur 1-year prior to failure. This treatment is commonly used by other researchers (Balcaen & Ooghe, 2006). Thus, the last available financial statements are used for one-year $(\mathrm{t}-1)$ prediction of the financial condition of firms.

For each firm, the four financial ratios from Altman's z"-score model are calculated: working capital to total assets (WCTA); retained earnings to total assets (RETA); earnings before interest and tax to total assets (EBITTA); book value of equity to total debt (BVETD). All financial ratios are used in their decimal format. The size of a firm is measured by the natural logarithm of its total assets (R '000s), similar to Altman et al. (2017). The age of a firm is taken as the number of years since inception, not the date each firm listed on the JSE. The date of inception was extracted from each firm's registration number.

Macroeconomic data for annual economic growth, annual inflation, and average annual lending rate is incorporated. Annual economic growth rate is taken as the percentage annual change in GDP such that:

$$\Delta\mathrm{GDP}_t = \frac{\mathrm{GDP}_t - \mathrm{GDP}_{t-1}}{\mathrm{GDP}_{t-1}} \times 100$$

Annual inflation is proxied by the percentage annual change in CPI such that:

$$\Delta\mathrm{CPI}_t = \frac{\mathrm{CPI}_t - CPI_{t-1}}{\mathrm{CPI}_{t-1}} \times 100$$

The prime lending rate (PLR) is the average annual lending rate. The macroeconomic variables are matched to the analysis year in which the accounting ratios are computed. This treatment of macroeconomic variables allows for variability of these factors within the sample as firms failed at various times in the period under consideration.

To test the research hypothesis, a benchmark model is required. This study considers the four variable logit model of Altman et al. (2017) as the benchmark:

$$L = 0.035 - 0.495x_1 - 0.862x_2 - 1.721x_3 - 0.017x_4$$

The logistic regression analysis is defined such that the dependent variable $Y = 0$ for non-failed firms and $Y = 1$ for failed firms. The accounting variables are: $x_1$ is Working Capital to Total Assets (WCTA); $x_2$ is Retained Earnings to Total Assets (RETA); $x_3$ is EBIT to Total Assets (EBITTA); and $x_4$ is Book Value of Equity to Total Debt (BVETD). The classification performance of this model is assessed by the AUC measure.

The first hypothesis (H1) is tested by re-estimating the coefficients of the benchmark model's coefficients using the South African firm data of this study. The resulting logit is named Model 1.

$$L = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$$

Similar to the results of Altman et al. (2017), we expect the coefficients of the accounting ratios to have negative signs. This is consistent with the expectation that an increase in these ratios would result in a reduced probability of failure. The classification performance of Model 1 is assessed by the AUC measure. Hypothesis H1 will be supported if the AUC of Model 1 exceeds the AUC of the benchmark model.

The second hypothesis (H2) is tested by including an additional variable for size, which is measured by the natural logarithm of total assets, to Model 1. It is important to note that total assets data is in thousands (R '000s). Thus, Model 2 for testing this hypothesis has the following logit form:

$$L = b_0 + \sum_{i=1}^{4} b_ix_i + b_5\ln(TA)$$

In Gretl, the natural logarithm of total assets variable is labelled "LNTA" for simplicity. The coefficient for the firm size variable, $b_5$, is expected to be negative as larger firms have a lower probability of failure. Hypothesis H2 will be supported if the AUC of Model 2 exceeds the AUC of the benchmark model.

The third hypothesis (H3) is tested by including firm age as an additional variable to the four financial ratios used in Model 1. Firm age is measure in years since inception. The resulting Model 3 has a logit with the form:

$$L = b_0 + \sum_{i=1}^{4} b_ix_i + b_6AGE$$

The coefficient for firm age, $b_6$, is expected to be negative indicating that older firms have a lower probability of failure. If the AUC of Model 3 is higher than that of the benchmark model, then our hypothesis H3 is supported.

The forth hypothesis (H4) is tested by including the three macroeconomic variables to Model 1. The resulting Model 4 has a logit in the form of:

$$L = b_0 + \sum_{i=1}^{4} b_ix_i + b_7\Delta GDP + b_8\Delta CPI + b_9PLR$$

Lower GDP growth rate or recessions tend to increase the probability of firm failure; thus, we expect its coefficient, $b_7$, to be negative. An increase in inflation or interest rates is expected to have a negative impact on firm financial health as labour costs and cost of debt increases; thus, the coefficients $b_8$ and $b_9$ are expected to be positive. Hypothesis H4 is supported if the AUC of Model 4 exceeds the AUC of the benchmark model.

The fifth hypothesis (H6) is tested by including the three super sector dummy variables to Model 1. The resulting Model 5 has a logit in the form of:

$$L = b_0 + \sum_{i=1}^{4} b_i x_i + \sum_{j=1}^{3} c_j D_j$$

The JSE classifies firms based on the Industry Classification Benchmark (ICB) (FTSE, 2017) system developed by FTSE Russel. The structure of the ICB is to categorise companies into 10 industries which are further categorised into super sectors. This study will focus on the super sector classification due to the limited number of firms. An examination of failed firms in-sample revealed that three super sectors were frequent. There were 6 failed firms in the construction & materials (ICB code: 2300) super sector; 7 failed firms in the industrial goods & services (ICB code: 2700) super sector; and 11 failed firms in the technology (ICB code: 9500) super sector. These three super sectors are included in the Logit model as dummy variables where: $D_1$ = 1 when a firm is in the construction & materials super sector, 0 otherwise; $D_2$ = 1 when a firm is in the industrial goods & services super sector, 0 otherwise; $D_3$ = 1 when a firm is in the technology super sector, 0 otherwise. The coefficients for the three dummy variables are expected to be positive as a firm in these super sectors could be at a higher risk of failure. If the AUC of Model 5 is higher than the AUC of the benchmark model, then our hypothesis H5 is supported.

To determine the performance of all variables, Model 6 is studied with the following logit form:

$$L = b_0 + \sum_{i=1}^{4} b_i x_i + b_5 \ln(TA) + b_6 AGE + b_7 \Delta GDP + b_8 \Delta CPI + b_9 PLR + \sum_{j=1}^{3} c_j D_j$$

If the AUC of Model 6 is higher than the AUC of the benchmark model, then the inclusion of firm size, firm age, macroeconomic variables, and super sector dummy variables will be supported.

In the case that some variables in Model 6 are not statistically significant at the 5% level, an improved model is investigated. This model is estimated by sequentially eliminating variables not statistically significant at the 5% level using a two sided test (Gujarati & Porter, 2009). The resulting model with only statistically significant variables is Model 7.

## 3.7. Econometric software

Gretl software was used to perform all econometric analysis in this report. Gretl is a free software that is flexible, extensible, and accurate which makes it a suitable alternative to commercial software. The details of the Gretl and extensions used in this study is as follows:

- Gretl 2017c
- ROC package (version 1.02) for the computation and analysis of ROC for binary choice models
- lp-mfx package (version 0.4) for computation of logit and probit marginal effects

## 3.8. Summary

The research methodology has been presented in this chapter. The research hypotheses were clearly defined to guide the data analysis. The logit model was developed to test the adequacy of financial ratios, firm size, firm age, macroeconomic conditions, and industrial sectors in predicting corporate failure in South Africa. Yearly financial statements and time series economic indicator data was collected from various sources for the 16-year period from 2000 to 2015. Seven logit models will be developed using this in-sample data from 2000 to 2015. The developed models will be tested for performance on an out-of-sample set of firms to determine the accuracy of each model in predicting one-year financial failure or non-failure.

# Chapter 4 - Analysis and Results

## 4.1. Introduction

Results of econometric data analysis based on the methodology described in Chapter 3 are presented in this section. These results are for one-year prior to failure prediction for in-sample data and out-of-sample data. For in-sample data analysis, statistics for failed and non-failed data are presented for the 2000 to 2015 period in this study. Seven Logit models are developed from the analysis of in-sample data. The developed models are analysed for coefficient signs, marginal effects of each variable on the dichotomous outcome, and performance accuracy of each model using the AUC measure. The seven models are tested on the out-of-sample data to determine performance accuracy.

## 4.2. In-sample analysis

### 4.2.1. Summary statistics

The descriptive statistics of the four accounting ratios, firm size and firm age are presented for the entire data set in Table 8. Descriptive statistics for failed firms and for non-failed firms are discussed late in this section. The full sample data consists of failed firms with total assets ranging from R2.5 million to R1.3 billion, while non-failed firms range from R12.8 million to R15.1 billion in total assets.

*Table 8: Entire data summary statistics for one-year failure prediction*

| ***Entire data set*** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Mean | Minimum | Maximum | Std. Dev. | Skewness | Excess kurtosis | JB-stat | JB *p*-value |
| WCTA | 0.143 | -1.159 | 0.656 | 0.267 | -1.744 | 5.849 | 191.295 | 0.000 |
| RETA | -0.065 | -6.368 | 0.769 | 1.013 | -3.830 | 17.401 | 1491.12 | 0.000 |
| EBITTA | -0.001 | -4.026 | 0.590 | 0.450 | -7.380 | 63.268 | 17410.4 | 0.000 |
| BVETD | 1.224 | -0.567 | 11.925 | 1.588 | 4.145 | 22.486 | 2369.15 | 0.000 |
| LNTA | 13.235 | 7.812 | 16.533 | 1.910 | -0.369 | -0.113 | 2.3039 | 0.316 |
| AGE | 27.101 | 2.000 | 102.000 | 25.327 | 1.206 | 0.429 | 24.7532 | 0.000 |
| *Note*: Number of observations = 99. WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception. Jarque-Bera (JB) statistic with corresponding *p*-value | | | | | | | | |

The average age of firms in the sample is 27 years with the oldest analysed firm being 102 years old. Table 8 includes the Jarque-Bera (JB) test for normality with the null hypothesis that the variables are normally distributed with the corresponding $p$-value. The $p$-value represents the lowest significance level at which the null hypothesis of normality can be rejected. In this study, the acceptable level of significance, $\alpha$, is 5%. For all sample variables, the null hypothesis of normality is rejected based on p-values less than 0.001, except for LNTA for which we cannot reject the null hypothesis due to a p-value greater than 0.05.

The means accounting ratios of failed firms, shown in Table 9, are negative for $x_1$(WCTA), $x_2$ (RETA), and $x_3$ (EBITTA). This indicates that average failed firms have: poor net liquid assets management; have low cumulative earnings to reinvest; and poor asset productivity. These three accounting ratios have distributions which are negatively skewed with positive excess kurtosis. Some of the failed firms in the sample had negative $x_4$ (BVETD), observed by the minimum of -0.567, which indicates that prior to failure, total liabilities exceeded total assets held by some failed firms. The accounting ratio $x_4$ (BVETD) has a positively skewed distribution with negative excess kurtosis. Failed firms have a wide range of leverage as depicted by the high standard deviation for RETA. The size of failed firms, as measured using LNTA, has a distribution with a minor negative skewness and an excess kurtosis of 0 which indicates a mesokurtic distribution that is induced by the logarithmic transformation. Taking the logarithm of a variable reduces both skewness and excess kurtosis (Brooks, 2014). The average age of failed firms in the sample is 17.2 years with the oldest being 61 years old.

*Table 9: Failed firm summary statistics for one-year failure prediction*

| ***Failed Firms*** | | | | | | |
|---|---|---|---|---|---|---|
| Variable | Mean | Minimum | Maximum | Std. Dev. | Skewness | Ex. kurtosis |
| WCTA | -0.038 | -1.159 | 0.495 | 0.336 | -1.358 | 2.085 |
| RETA | -0.733 | -6.368 | 0.413 | 1.478 | -2.276 | 5.096 |
| EBITTA | -0.223 | -4.026 | 0.590 | 0.714 | -4.567 | 22.184 |
| BVETD | 0.557 | -0.567 | 2.339 | 0.684 | 0.585 | -0.321 |
| LNTA | 11.923 | 7.812 | 14.056 | 1.590 | -0.729 | 0.006 |
| AGE | 17.235 | 2.000 | 61.000 | 16.218 | 1.476 | 1.422 |

*Note*: Number of observations = 34. WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception

The means for non-failed or healthy firms, presented in Table 10, are positive for $x_1$(WCTA), $x_2$ (RETA), $x_3$ (EBITTA), and $x_4$ (BVETD). Positive mean values indicate that the average healthy firm has positive net liquid assets, finance assets through retained earnings, assets are productive, and total assets exceed total liabilities. The WCTA, BVETD and AGE of non-failed firms have positively skewed distributions. The RETA, EBITTA and LNTA exhibit negatively skewed distributions. The four financial ratios have positive excess kurtosis while size (LNTA) and age of non-failed firms have negative excess kurtosis. The average size and age of non-failed firms in the sample are, as expected, higher than for failed firms.

*Table 10: Non-failed firm summary statistics for one-year failure prediction*

| ***Non-failed firms*** | | | | | | |
|---|---|---|---|---|---|---|
| Variable | Mean | Minimum | Maximum | Std. Dev. | Skewness | Ex. kurtosis |
| WCTA | 0.238 | -0.006 | 0.656 | 0.155 | 0.805 | 0.048 |
| RETA | 0.285 | -0.716 | 0.769 | 0.287 | -1.579 | 3.671 |
| EBITTA | 0.115 | -0.261 | 0.363 | 0.083 | -0.975 | 5.865 |
| BVETD | 1.572 | 0.076 | 11.925 | 1.805 | 3.830 | 17.182 |
| LNTA | 13.921 | 9.460 | 16.533 | 1.700 | -0.487 | -0.093 |
| AGE | 32.262 | 3.000 | 102.000 | 27.711 | 0.907 | -0.401 |

*Note*: Number of observations = 65. WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception

### 4.2.2. Correlation matrix for all data

Table 11 presents the correlation matrix and multicollinearity for the four accounting ratios, firm size, firm age and macroeconomic variables for all data. The correlations between variables are low with two exceptions. The highest correlations were between WCTA and RETA at 0.602 and between ΔCPI and PLR at 0.554. However, all variables have VIF values well below 10 which indicates that the data set has no multicollinearity problems.

*Table 11: Correlation matrix and VIF for one-year failure prediction*

| | WCTA | RETA | EBITTA | BVETD | LNTA | AGE | ΔGDP | ΔCPI | PLR |
|---|---|---|---|---|---|---|---|---|---|
| WCTA | 1.000 | | | | | | | | |
| RETA | 0.602 | 1.000 | | | | | | | |
| EBITTA | 0.357 | 0.244 | 1.000 | | | | | | |
| BVETD | 0.499 | 0.222 | 0.142 | 1.000 | | | | | |
| LNTA | 0.260 | 0.467 | 0.385 | 0.062 | 1.000 | | | | |
| AGE | 0.174 | 0.265 | 0.154 | -0.036 | 0.370 | 1.000 | | | |
| ΔGDP | 0.033 | 0.082 | 0.064 | -0.098 | -0.201 | 0.035 | 1.000 | | |
| ΔCPI | -0.177 | -0.096 | -0.035 | -0.025 | -0.072 | 0.090 | -0.178 | 1.000 | |
| PLR | -0.167 | -0.069 | -0.107 | -0.110 | -0.338 | 0.137 | 0.430 | 0.554 | 1.000 |
| ***Multicollinearity*** | | | | | | | | | |
| VIF | 2.292 | 2.015 | 1.342 | 1.399 | 1.997 | 1.312 | 1.861 | 2.169 | 2.928 |

*Variables*: WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception; ΔGDP is economic growth rate; ΔCPI is inflation rate; PLR is prime lending rate; VIF is the Variance Inflation
*Note*: Multicollinearity is measured the VIF for which a value above 10 indicates strong linear dependency between two or more independent variables

### 4.2.3. Benchmark model results

The four-accounting ratio logit model of Altman et al. (2017), which is taken as the benchmark model, was applied to the sample data to predict failure within one-year, for South African firms. The results of the analysis are presented in Figure 2 as a ROC curve for the benchmark model's performance. The area under the ROC curve (AUC) is a measure of a model's discriminating performance between failed and non-failed firms. The AUC of the blue line in Figure 2 indicates a model with an AUC of 0.5 which is no better than a coin toss. Such a model is only accurate 50% of the time. The ROC of the benchmark model is depicted by the red line in Figure 2. A perfect model would have an ROC curve along the left vertical axis and top horizontal axis, indicating a True Positive Rate of 1 for all False Positive Rates. The computed AUC for the benchmark model is 0.90 with a standard error of 0.038. The 95% confidence interval for the benchmark model's AUC is [0.826, 0.973]. Altman et al. (2017) found the four-accounting variable model to have an AUC of 0748 for all data in their study with only one country having an AUC above 0.90. This confirms that the Altman et al. (2017) four-accounting variable model is applicable to South African firms.

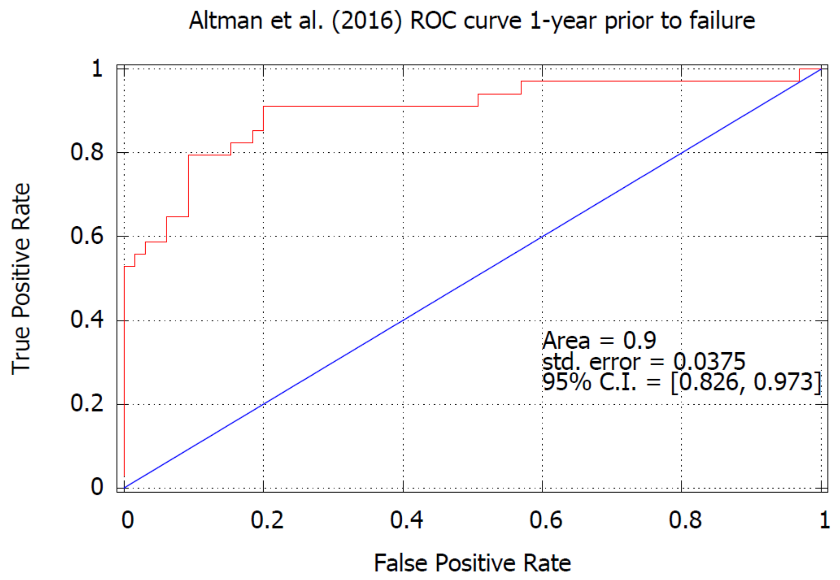


*Figure 2: Altman et al. (2017) ROC curve for 1-year prior to failure prediction*

### 4.2.4. Logit analysis results

Seven models were analysed, as discussed in the research methodology, to test the postulated hypotheses. Model 1 represents the re-estimated coefficients of the four accounting ratios based on this study's sample data. Model 2 represents the addition of firm size to the four accounting ratios model. Model 3 explicitly considers the age of the firm along with the four accounting ratios. Model 4 incorporated three macroeconomic variables to the four accounting ratios model. Model 5 considers sector effects by including three dummy variables to the four accounting ratios. The dummy variables aim to capture the propensity of failure for firms in the construction & materials, industrial goods & services, and technology sectors. Model 6 constitutes all variables considered in this study: accounting ratio; firm size; firm age; macroeconomic variables; and sector. The parsimonious model, Model 7, contains only those variables that are statistically significant in explaining the probability of a firm failing. The logistic regression analysis results of these models are presented in Table 12. For ease of reference, the benchmark model coefficients for the four accounting variables are include in Table 12.

*Table 12: Logistic regression model results for one-year failure prediction*

| Variable | Benchmark | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
|---|---|---|---|---|---|---|---|---|
| Constant | 0.035 | 0.959<br>(1.429) | 8.770**<br>(2.769) | 2.054**<br>(2.000) | -4.408**<br>(-2.299) | 0.581<br>(0.722) | -0.779<br>(-0.115) | -6.888**<br>(-2.497) |
| WCTA | -0.495 | -5.440*<br>(-1.831) | -8.009**<br>(-2.252) | -6.721*<br>(-1.938) | -6.256*<br>(-1.670) | -5.885*<br>(-1.918) | -14.269**<br>(-1.972) | -13.218**<br>(-2.146) |
| RETA | -0.862 | -2.267**<br>(-2.115) | -0.870<br>(-0.791) | -1.777<br>(-1.617) | -2.981**<br>(-2.530) | -2.058*<br>(-1.830) | -1.734<br>(-1.072) | -2.807**<br>(-2.058) |
| EBITTA | -1.721 | -6.609**<br>(-2.589) | -6.660**<br>(-2.554) | -7.122**<br>(-2.723) | -6.535**<br>(-2.507) | -6.062**<br>(-2.383) | -9.803**<br>(-2.334) | -10.057**<br>(-2.631) |
| BVETD | -0.017 | -0.367<br>(-0.740) | -0.427<br>(-0.831) | -0.606<br>(-1.057) | -0.150<br>(-0.417) | -0.291<br>(-0.559) | -0.105<br>(-0.164) | |
| LNTA | | | -0.566**<br>(-2.573) | | | | -0.419<br>(-1.077) | |
| AGE | | | | -0.028*<br>(-1.662) | | | -0.051*<br>(-1.700) | -0.055**<br>(-2.116) |
| ΔGDP | | | | | -0.344<br>(-1.148) | | -0.891*<br>(-1.887) | -0.802**<br>(-1.993) |
| ΔCPI | | | | | -0.289<br>(-1.271) | | -0.736*<br>(-1.864) | -0.758**<br>(-2.060) |
| PLR | | | | | 0.711**<br>(2.797) | | 1.410**<br>(2.786) | 1.486**<br>(3.144) |
| $D_1$ | | | | | | 0.869<br>(0.912) | 3.515**<br>(2.232) | 2.915**<br>(2.310) |
| $D_2$ | | | | | | 0.041<br>(0.049) | 0.770<br>(0.597) | |
| $D_3$ | | | | | | 0.920<br>(0.977) | 0.943<br>(0.636) | |
| n | - | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| McFadden $R^2$ | - | 0.486 | 0.546 | 0.513 | 0.587 | 0.499 | 0.704 | 0.692 |
| Log-Likelihood | - | -32.7 | -28.9 | -31 | -26.3 | -31.9 | -18.8 | -19.6 |

*Variables*: Dependent variable Y = 1 for failed, 0 for non-failed. WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception; ΔGDP is economic growth rate; ΔCPI is inflation rate; PLR is prime lending rate; $D_1$ is a dummy variable for the construction & materials super sector; $D_2$ is a dummy variable for the industrial goods & services super sector; and $D_3$ is a dummy variable for the technology super sector

*Models*: Benchmark = Altman et al. (2017) four variable logit model; Model 1 = Accounting ratio logit model; Model 2 = Logit model estimated with firm size variable; Model 3 = Logit model estimated with firm age variable; Model 4 = Logit model estimated with macroeconomic variables; Model 5 = Logit model estimated with industry/sector dummy variables; Model 6 = Logit model estimated with all variables; Model 7 = Logit model containing only statistically significant variables.

*Notes*: * indicates significance at the 10% level. ** indicates significance at the 5% level. The t-statistics are given in parentheses for all coefficients.

The signs of the coefficients presented in Table 12 may be interpreted to indicate the influence of each variable on the probability of failure, however, the magnitudes of the coefficients cannot be interpreted as elasticity as in Ordinary Least Squares (OLS). In conditional probability models, such as the Logit model, marginal effects are computed to allow interpretation of how the variables affect the probability of failure. The marginal effects of the dependent variables for the seven logit models are presented in Table 13.

*Table 13: Logit model marginal effects*

| ***Marginal effects*** *($\partial P/\partial x_i$)* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Variable | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 |
| WCTA | -1.356 | -1.949 | -1.656 | -1.555 | -1.456 | -3.490 | -3.296 |
| RETA | -0.565 | -0.212 | -0.425 | -0.741 | -0.509 | -0.424 | -0.700 |
| EBITTA | -1.647 | -1.621 | -1.742 | -1.624 | -1.500 | -2.397 | -2.508 |
| BVETD | -0.091 | -0.104 | -0.147 | -0.037 | -0.072 | -0.026 | |
| LNTA | | -0.138 | | | | -0.102 | |
| AGE | | | -0.007 | | | -0.012 | -0.014 |
| ΔGDP | | | | -0.085 | | -0.218 | -0.200 |
| ΔCPI | | | | -0.072 | | -0.180 | -0.189 |
| PLR | | | | 0.177 | | 0.345 | 0.371 |
| $D_1$ | | | | | 0.213 | 0.621 | 0.538 |
| $D_2$ | | | | | 0.010 | 0.189 | |
| $D_3$ | | | | | 0.225 | 0.231 | |

*Variables*: Dependent variable Y = 1 for failed, 0 for non-failed. WCTA is Working Capital/Total Assets; RETA is Retained Earnings/Total Assets; EBITTA is EBIT/Total Assets; BVETD is Book Value of Equity/Total Debt; LNTA is the natural logarithm of Total Assets; AGE is the number of years since inception; ΔGDP is economic growth rate; ΔCPI is inflation rate; PLR is prime lending rate; $D_1$ is a dummy variable for the construction & materials super sector; $D_2$ is a dummy variable for the industrial goods & services super sector; and $D_3$ is a dummy variable for the technology super sector

*Models*: Model 1 = Accounting ratio logit model; Model 2 = Logit model estimated with firm size variable; Model 3 = Logit model estimated with firm age variable; Model 4 = Logit model estimated with macroeconomic variables; Model 5 = Logit model estimated with industry/sector dummy variables; Model 6 = Logit model estimated with all variables; Model 7 = Logit model containing only statistically significant variables. $D_1$ is a dummy variable for the construction & materials super sector; $D_2$ is a dummy variable for the industrial goods & services super sector; and $D_3$ is a dummy variable for the technology super sector.

Model 1 results show that the four accounting ratios coefficients, in Table 12, are all negative, consistent with the benchmark model. Negative coefficients are consistent with economic theory and expectations that an increase in any of the accounting ratios of a firm will decrease its probability of failure. An increase in any accounting variable with a negative coefficient causes the logit to be increasingly negative which results in a lower probability of failure. The WCTA, RETA, and EBITTA accounting variables are statistically significant at the 10% level. BVETD is not statistically significant for the sample data analysed. This model's McFadden $R^2$ of 0.486 and log-likelihood of -32.7 will be used for comparison with the other models. From Table 13, WCTA and EBITTA have the highest marginal effects on the probability of failure, while BVETD has the least effect. To determine the effect of individual variables, an increase in a variable multiplied by its marginal effect gives the expected change in the probability of failure, with all other variables held constant. The AUC of Model 1 is 0.907, as determined from the ROC curve shown in Figure 3, is a measure of the model's discriminating accuracy. The AUC is higher than the benchmark model (AUC = 0.902) which supports the hypothesis H1. The re-estimation of the coefficients with South African firm data results in marginal accuracy improvement of the benchmark model.

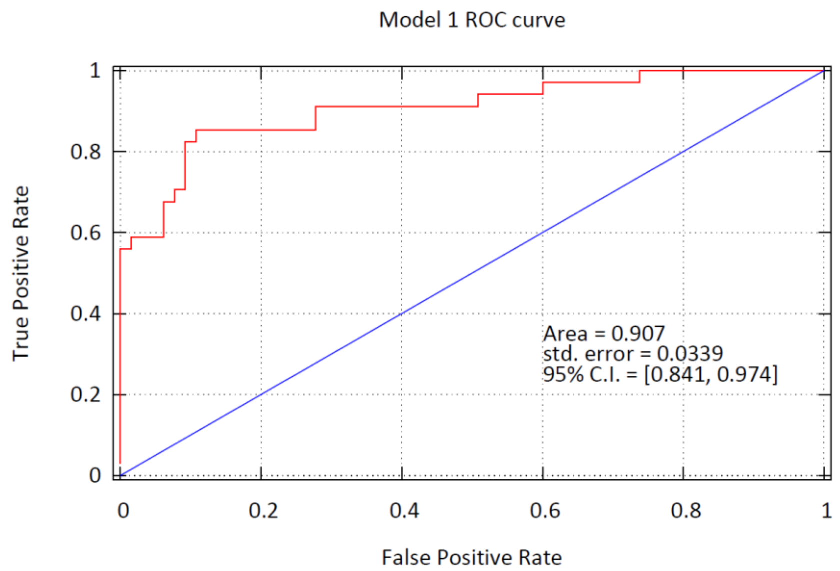


*Figure 3: Model 1 ROC curve for 1-year prior to failure*

Results for Model 2, which includes a variable for firm size, in Table 12 show that the four accounting variables have negative coefficients as expected. The coefficient for the firm size variable, LNTA, is negative which confirms that larger firms have a lower probability of failure than smaller firms. The WCTA and EBITTA accounting variables and LNTA are statistically significant at the 5% level. When firm size is explicitly considered, the RETA and BVETD accounting ratios are statistically not significant in explaining probability of failure. A McFadden $R^2$ of 0.546 indicates better goodness of fit compared to Model 1, thus it fits the data better. From Table 13, WCTA and EBITTA have the highest marginal effects on the predicted financial condition of a firm. The marginal effect of RETA, compared to Model 1, is reduced as the accounting ratio is statistically not significant. A unit increase in LNTA would reduce the probability of failure by 13.8%, if the other variables are held constant. The ROC curve for Model 2, shown in Figure 4, has an AUC of 0.93 which is an improvement on the performance accuracy of the benchmark model (AUC = 0.90). The inclusion of firm size improves the accuracy of the benchmark model and Model 1, which supports the hypothesis H2.

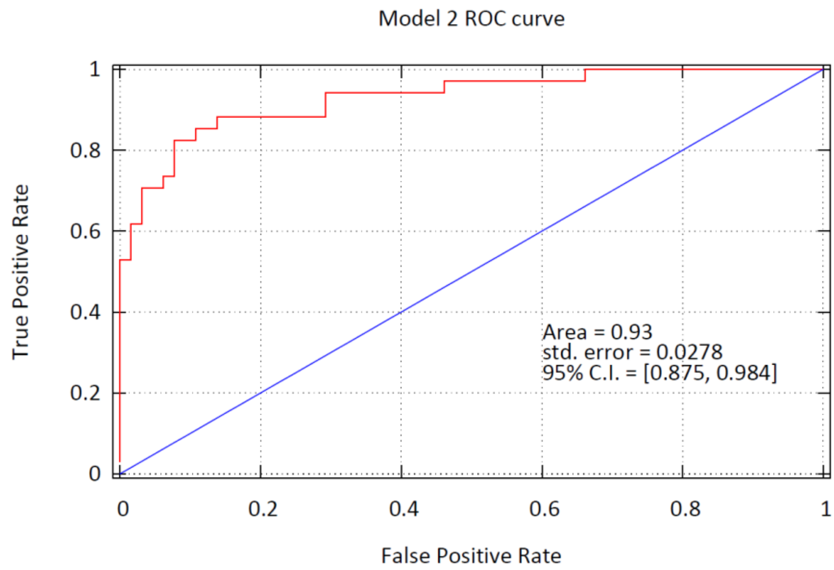


*Figure 4: Model 2 ROC curve for 1-year prior to failure*

Coefficients for the four accounting variables of Model 3, in Table 12, are all negative as expected. The coefficient for AGE is negative giving support to the notion of older firms having a lower risk of failure. The variables statistically significant at the 10% level for Model 3 are WCTA, EBITTA and AGE. RETA and BVETD are statistically not significant when firm age is explicitly considered in a prediction model. The log likelihood function is maximised more than Model 1 but lower than Model 2 which suggests that Model 3 is a marginal improvement on Model 1. WCTA and EBITTA have the highest marginal effects on the prediction outcome, based on results in Table 13. The marginal effect of AGE suggests that a unit increase in the age of a firm reduces the probability of failure by 0.7%, when other variables are held constant. Figure 5 shows the ROC curve for Model 3 with an AUC of 0.909 which is an improvement on the discriminating accuracy of the benchmark model (AUC = 0.90). This result is consistent with the log likelihood result in that Model 3 is a marginal improvement on Model 1 with an AUC of 0.907. The explicit inclusion of a firm’s age improves the accuracy of the benchmark mode, which supports the hypothesis H3.

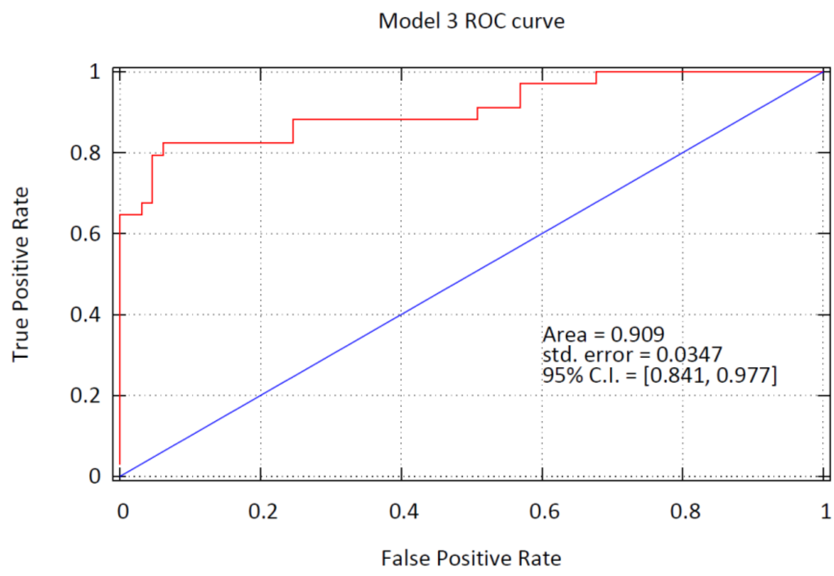


*Figure 5: Model 3 ROC curve for 1-year prior to failure*

The results for Model 4, in Table 12, show that all accounting variables have negative coefficients as expected. The coefficient for GDP is negative, thus an increasing in economic activity reduces firm failure probability. The coefficient for PLR is positive, confirming that higher lending rates increase the probability of firm failure. The coefficient for CPI is negative, which indicates that higher inflation reduces the probability of firm failure. Such a scenario is possible when inflation is driven by aggregate demand, rather than aggregate supply. An increase in aggregate demand can be the result of various macroeconomic dynamics such as increased government or consumer spending. In the short run, inflation can increase due to aggregate demand in the economy increasing faster than the increase in aggregate supply. Results show that all variables in the model are statistically significant at the 10% level except for BVETD, ΔGDP, and ΔCPI. The model maximises the log likelihood function to -26.3 which is an improvement on Model 1. The effect of macroeconomic variable changes is assessed based on the marginal effects in Table 13. A unit percentage increase in the GDP or economic growth rate and inflation reduces the probability of failure by 8.5% and 7.2%, respectively. A unit percentage increase in the PLR increases the chances of a firm failing with a year by 17.7%. The AUC of Model 4 is 0.943, computed from the ROC curve in Figure 6, which is higher than the benchmark model (AUC = 0.90). The performance accuracy measured by the AUC supports the hypothesis H4 that macroeconomic variables improve the accuracy of the benchmark model.

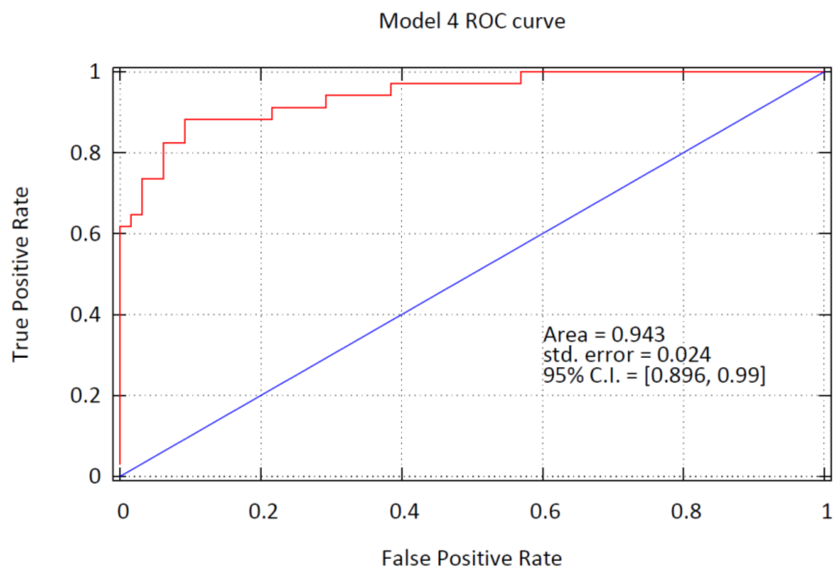


*Figure 6: Model 4 ROC curve for 1-year prior to failure*

Results in Table 12 for Model 5 indicate that coefficients of all accounting variables are positive. The coefficients for the dummy variables representing the construction & materials, industrial goods & services, and technology super sectors of the JSE are all positive. These results reaffirm that a firm in any of these three super sectors has a higher probability of failure than other super sectors. The variables that remain statistically significant within the acceptable level of 10% are WCTA, EBITTA, LNTA. BVETD and the three dummy variables are statistically not significant in this model. The model has a log likelihood value of -31.9 which is marginal improvement to Model 1. Table 13 results indicate that WCTA and EBITTA have the highest marginal effect on the prediction outcome. The effect of super sectors may be assessed based on the marginal effects of the three dummy variables. A firm that operates in the construction & materials super sector has a 21.3% risk of failure. A firm in the industrial goods & services super sector has a 10% risk of failure. A firm in the technology super sector has a 22.5% probability of failure. Model 5 has an AUC of 0.911, computed from the ROC curve in Figure 7, which is higher than the benchmark model (AUC = 0.90). These results support that a model with sector dummy variables will result in higher performance accuracy compared to the benchmark model, thereby supporting the hypothesis H5.

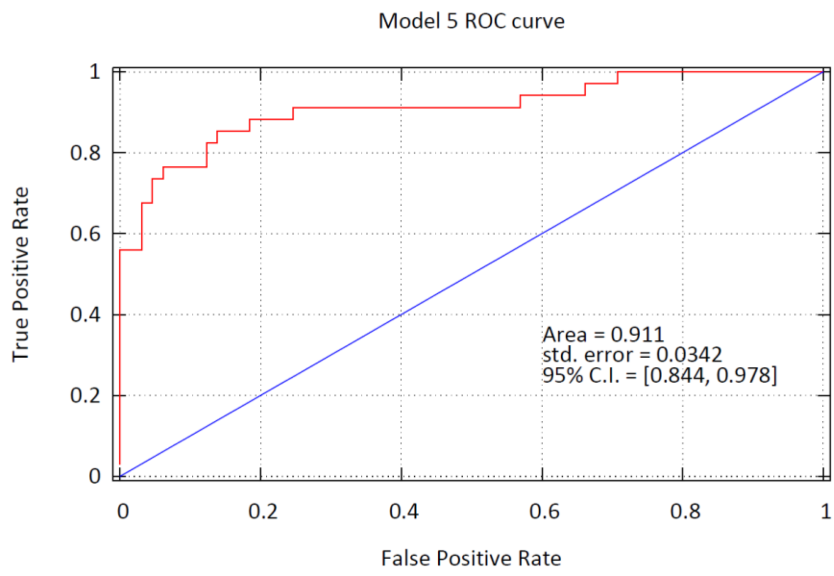


*Figure 7: Model 5 ROC curve for 1-year prior to failure*

Model 6 which includes accounting, firm size, firm age, macroeconomic and sector variables to predict failure is presented in Table 12. As with the other five models, the coefficients of the four accounting variables are negative. The coefficients for LNTA and AGE are negative which is consistent with the results from Model 2 and Model 3, respectively. The coefficients for ΔGDP and ΔCPI are negative, while PLR has a positive coefficient which is consistent with Model 4 results. The three super sector dummy variables have positive coefficients as found in Model 5. When all variables are included in a model, their influence on the probability of a firm failing is consistent with previous models 1 to 5. The WCTA and EBITTA accounting variables, AGE variable, three macroeconomic variables, and $D_1$ dummy variable are statistically significant at the 10% level. The variables that are statistically not significant in predicting failure are: RETA; BVETD; LNTA; $D_2$; and $D_3$. This model has a log likelihood value of -18.8 which the highest of all the models considered thus far. As this model best maximises the log likelihood function, it is expected to offer the highest prediction accuracy compared to the previous five models. From the results in Table 13, the marginal effect of each variable can be assessed when all variables are incorporated into one model. The WCTA and EBITTA accounting variables have the highest effect on the probability of failure. A unit increase in firm size by measure of LNTA decreases the probability of failure by 10.2%. When all variables are considered, a unit increase in firm AGE reduces the probability of failure by 1.2%. Unit percentage increases in economic growth rate and inflation reduces the probability of failure by 21.8% and 18%, respectively. However, a unit increase in the prime lending rate, PLR, increases the risk of failure by 34.5%. When all variables are accounted for, firms in the construction & materials super sector are 62.1% more likely to fail. This significant increase in risk compared to Model 5 could be the result of inclusion of firm size, firm age, and macroeconomic variables which highlight the uncertainty in the construction & materials sector.

The predictive performance of the model is shown in the ROC curve, Figure 8. The performance accuracy of the model is an AUC of 0.972 which is higher than the benchmark model (AUC = 0.90) and other models studied in this research.

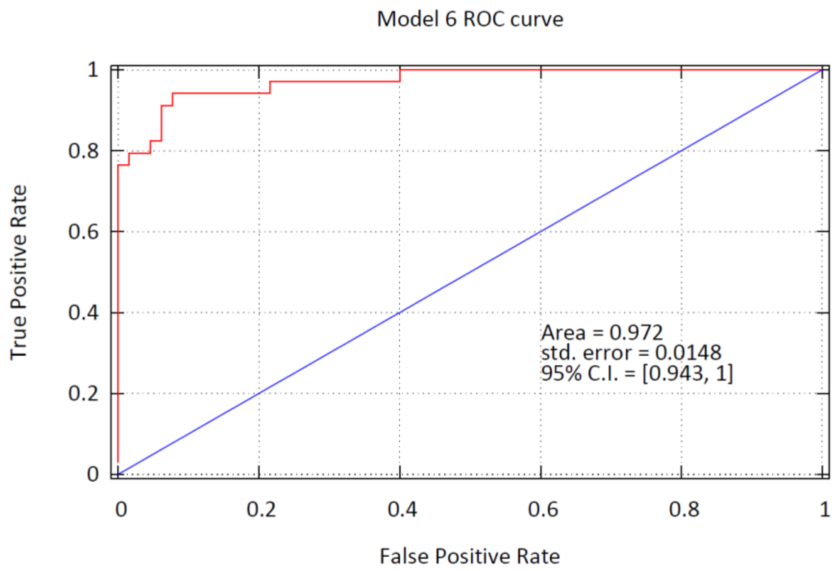


*Figure 8: Model 6 ROC curve for 1-year prior to failure*

A parsimonious model was developed by taking the results of Model 6 and sequentially eliminating variables with *p*-values higher than 0.05. The elimination of variables stops when a model with only statistically significant variables at the 5% level is achieved. This Model 7 is presented in Table 12. The final parsimonious model consists of three accounting variables, firm age and three macroeconomic variables and a dummy variable for the construction & materials super sector. The accounting variables that are statistically significant in failure prediction are: WCTA which measures a firm's liquid assets relative to its total debt and equity; RETA which measures the leverage of a firm; and EBITTA which measures the productivity of a firm's assets independent of any leverage or tax considerations. WCTA has the highest marginal effect of the three accounting variables which highlights the importance of liquid assets in reducing failure probability. A unit increase in a firm's age reducing the risk of failure by 1.4%. The effect of explicit firm age may be reduced by the inclusion of RETA in the model as retained earnings, which measure cumulative profitability, implicitly consider the age of a firm. The marginal effect of ΔGDP is 20% for every unit percentage increase in the economic growth rate. A unit increase in aggregate supply induced inflation rate reduce firm failure risk by 18.9%. In Model 7, the influence of the prime lending rate, PLR, is more pronounced as a unit percentage increase will increase risk of failure by 37.1%. Keeping all factors constant, a firm in the construction & materials super sector is 53.8% more likely to fail than all other super sectors. Model 7 maximises the log likelihood to -19.6 which is lower than Model 6 with a log likelihood value of -18.8. This marginal difference in log likelihood values is reflected in the performance accuracy of Model 7, shown in Figure 9, with an AUC of 0.971 compared to Model 6 with an AUC of 0.972. From the AUC results, Model 7's performance is equivalent to the more elaborate Model 6 in discriminating failed and non-failed firms. These results validate the parsimonious model.

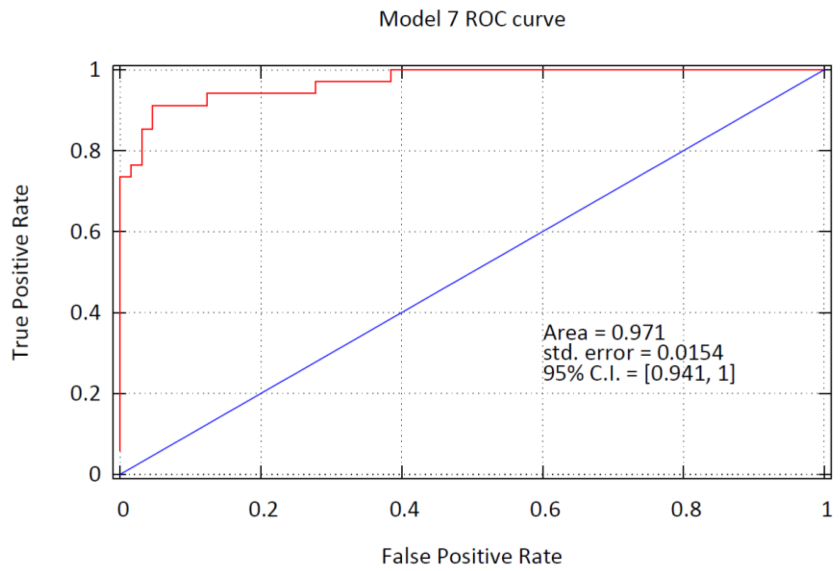


*Figure 9: Model 7 ROC curve for 1-year prior to failure*

The performance accuracies of the seven models are summarised in Table 14. For each model, Table 14, presents the AUC, 95% confidence interval for the AUC, and the standard error in computing the AUC. The best performing models, Model 6 and Model 7, have identical AUC standard errors. This result indicates that the statistically not significant variables in Model 7 only contribute 0.1% to the prediction accuracy. Based on this in-sample data analysis results, the parsimonious Model 7 is recommended for predicting failure of firms within a year.

*Table 14: Model prediction performance measures*

| Models | AUC | 95% Confidence Interval | AUC std. error |
|---|---|---|---|
| Benchmark | 0.900 | [0.826, 0.973] | 0.038 |
| Model 1 | 0.907 | [0.841, 0.974] | 0.034 |
| Model 2 | 0.930 | [0.875, 0.984] | 0.028 |
| Model 3 | 0.909 | [0.841, 0.977] | 0.035 |
| Model 4 | 0.943 | [0.896, 0.990] | 0.024 |
| Model 5 | 0.911 | [0.844, 0.978] | 0.034 |
| Model 6 | 0.972 | [0.943, 1] | 0.015 |
| Model 7 | 0.971 | [0.941, 1] | 0.015 |

## 4.3. Out-of-sample analysis

The models in Table 12 are tested on the out-of-sample data to determine the performance of each model in discriminating distressed and non-distressed firms. Financial statement data and macroeconomic factors from 2015 are used to predict the financial condition of the out-of-sample firms in 2016. Financial and macroeconomic data is used to calculate the value of the Logit, then the probability of distress for each firm. To determine the financial condition predicted by each model, the probability of distress cut-off $p = 0.5$ is used. Firms with $p < 0.5$ are predicted to be non-distressed or financially healthy and firms with $p \geq 0.5$ are predicted to be distressed. The results are presented in Table 15.

*Table 15: Correct predictions of models for out-of-sample distress prediction in 2016*

| Model | Non-failed | Failed |
|---|---|---|
| Model 1 | 90.91% | 100.00% |
| Model 2 | 100.00% | 100.00% |
| Model 3 | 90.91% | 100.00% |
| Model 4 | 100.00% | 100.00% |
| Model 5 | 100.00% | 83.33% |
| Model 6 | 90.91% | 100.00% |
| Model 7 | 81.82% | 100.00% |

Model 1 and Model 3 correctly predicted all distressed firms and 90.91% of the non-distressed firms. Both models incorrectly classified Curro Holdings Ltd as distressed when the company was in a healthy financial condition. Model 2 and Model 4 correctly classified all distressed firms and non-distressed firms. Model 5 correctly classified all non-distressed firms and 83% of distressed firms. The model incorrectly classified Imbalie Beauty as non-distressed with a probability of 0.46, when in fact the company is in financial decline. Model 6 correctly classified all distressed firms and 90.91% of the non-distressed firms. The model classified Accentuate Ltd as distressed when the company was in a healthy financial condition in 2016. Model 7 correctly predicted all distressed firms but only 81.8% of healthy firms. The model incorrectly predicted that Accentuate Ltd and Curro Holdings would be in financial decline in 2016 when the firms were in a healthy financial condition.

For out-of-sample failure prediction, based on the results in Table 15, the best performing models are those that contain the four accounting variables along with explicit inclusion of firm size (Model 2) or macroeconomic variables (Model 4). Model 6, which contains financial, firm size, firm age, macroeconomic and dummy variables, performed better than the parsimonious Model 7 in predicting the financial condition of non-failed firms. It is worth noting that all Models, except for Model 5 which incorporates super sector dummy variables, correctly predicted all firms in financial decline.

## 4.4. Summary

The sample data was analysed based on the descriptive statistics of failed, non-failed, and all firms. The statistics show that failed firms are on average smaller in size and younger than non-failed firms. All variables were checked to ensure that multicollinearity that could result in poor regression coefficients did not exist. Models were estimated using the logistic regression analysis and results presented as coefficients of the logit model's variables. Marginal effects were computed to determine the effect of individual variables on the probability of firm failure. The accuracy of each model was determined from ROC curves by computing the AUC. The AUC measures indicate that all developed models performed better than the benchmark model. These models were tested on an out-of-sample data set to determine their applicability in determine the recent financial condition of firms.

# Chapter 5 - Conclusion and Recommendations

## 5.1. Introduction

The principal objective of this study was to test the applicability of failure prediction models from developed economies on South African firms. The logit model of Altman et al. (2017), using only financial ratios, was applied to a sample of South African firms to determine its discriminating performance in classifying firms as failed or non-failed. The sample was taken over a 16-year period from 2000 to 2015 which included the global financial crisis. The secondary objective was to develop a logit model that accounts for the effect of financial ratios, structural factors and macroeconomic conditions in predicting firm failure. The resulting model performed better than the international model for in-sample and out-of-sample data.

## 5.2. Concluding comments

This study analysed the effect of financial, structural and macroeconomic variables on the failure of 99 JSE listed firms in South Africa during the period 2000 to 2015 using yearly cross-sectional data. To predict failure within one year, logistic regression analysis using financial ratios, firm size, firm age, macroeconomic variables, and industrial sectors as regressors was applied. Results were compared to the logistic regression model of Altman et al. (2017) which was developed using international private firm data from 28 European and 3 non-European countries.

When the logit model of Altman et al. (2017), using four accounting ratios as predictors, was applied to our sample data, its performance accuracy in discriminating failed and non-failed firms was 90%. This result confirms that the model, based on developed economies, is applicable to a developing economy such as South Africa. Re-estimating the coefficients (Model 1) using South African firm data marginally improved the accuracy of the international model. The addition of firm size (Model 2) and macroeconomic conditions (Model 4) has a strong effect on prediction accuracy. The improvement on model accuracy is marginal when firm age (Model 3) and industry super sectors (Model 5) are considered. Including all variables (Model 6) in the same model significantly increased the performance accuracy to 97.2% for in-sample data. However, a parsimonious model containing only statistically significant variables (Model 7) performed on par with the all variable model. This empirical study confirms the findings of Altman et al. (2017) that classification accuracy improves above 90% when using country-specific data with additional variables.

From the analytical results, it is possible to discuss the effect of individual variables on failure prediction models. The working capital to total assets (WCTA) and EBIT to total assets (EBITTA) financial ratios are strong predictors of failure, while the book value of equity to total debt (BVETD) is a weak predictor.  These results suggest that firms should maintain good liquidity and optimise the productivity of their assets.  Larger firms, measured by total assets, and older firms have a lower probability of failure.  Higher economic growth and inflation, driven by aggregate demand, are critical macroeconomic factors for firms to remain successful.  Higher lending rates, which result in higher cost of debt, have a detrimental effect on firms as a going concern.  Firms operating in the construction & materials super sector have a higher risk of failure, followed by the technology super sector.

Based on the in-sample and out-of-sample performance of all models, the model including all variables best discriminates between failed and non-failed firms in both samples.  Although the model contains variables which are statistically not significant, these variables carry valuable information about individual firms.

## 5.3. Recommendations

The results of this study are important to all stakeholders in a business, but more so to management and debtholders.  Management can apply these results internally to perform periodic assessment of the firm's risk of distress based on timeous financial data.  The focus should be on liquidity and operating profit before financing costs and taxes.  A balancing act is required to maintain a reasonable level of liquidity without reaching elevated levels of liquidity that are unprofitable.  Managers should monitor the productivity of the firm's assets which the goal of optimising productivity to realise maximum earnings from them.  While monitoring internal performance, managers should be aware of prevailing macroeconomic conditions and possible shocks that could introduce distress risk.  All internal factors remaining constant, high prime lending rates increase the risk of firm distress and possible failure.  High interest rates reduce a firm's liquidity which must be used to meet increasing borrowing costs.  Managers should be aware of this macroeconomic shock and plan to avoid financial difficulty.

Debtholders may use this research for screening before issuing new debt.  For debtholders such as banks, this is important to minimise the number of non-performing loans held by them.  For existing debt, this research indicates that debtholders should exercise caution with firms in the construction and services super sector.

The results of this study show, unequivocally, the critical impact of macroeconomic conditions on firm survival.  Policy makers in government, in setting monetary policy, should be aware of the relationship between firm failure and the macroeconomic environment.  The primary focus should be to stimulate economic growth.  When the economy is growing, on average, firms are also growing due to higher levels of aggregate demand which has a significant impact on the country's socio-economic environment. Growing economies attract foreign direct investment, reduce unemployment rates, and the survival rate of firms increases.

The South African Reserve Bank, in its monetary policy management adjusts the base lending rate, which is related to the prime lending rate, to keep inflation within a set target. Results from this study suggest that the Reserve Bank should target an inflation rate that is near the upper limit of the target range to reduce firm failures. An inflation rate that is at the lower limit of the target range will have the adverse effect of increasing the risk of firm failures. This is premised on inflation being driven by aggregate demand in the economy. When inflation is driven by aggregate supply, lower inflation rates will result in higher firm survival rates.

In setting the interest rate, the Reserve Bank has the most direct impact of firm failure probability. This link between the prime lending rates and firm failures should be taken into consideration when the Bank considers increasing rates. When the prime lending rate is high, firms reduce their liquidity to service the interest expense on debt which could lead to a liquidity crisis and financial distress. The impact could be more severe in sectors such as the construction and services sector where failure results in high numbers job losses.

Understanding the effect of the macroeconomic environment on the financial wellbeing of firms is important to all stakeholders such as employees, management, customers, suppliers, debtholders and government institutions. Application of the work presented in this study can offer early warning of potential financial distress of a firm within a year. Knowing this, interventions can be made to find immediate solutions to prevent the decline.

## 5.4. Further studies

This study developed a simple logit model to predict corporate failure within one year using financial, structural and macroeconomic data. Further research should focus on the following:

- Extending the prediction horizon of the model
- Testing the applicability of the models in other African countries and developing economies
- Developing a hazard model with the same variables
- Using genetic algorithms to determine the best financial ratios to use as predictors

# Chapter 6 - References


Acosta-González, E., & Fernández-Rodríguez, F. (2014). Forecasting Financial Failure of Firms via Genetic Algorithms. *Computational Economics, 43*(2), 133-157. doi:10.1007/s10614-013-9392-9

Acosta-González, E., Fernández-Rodríguez, F., & Ganga, H. (2017). Predicting Corporate Financial Failure Using Macroeconomic Variables and Accounting Data. *Computational Economics*. doi:10.1007/s10614-017-9737-x

Adnan Aziz, M., & Dar, H. A. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance: The international journal of business in society, 6*(1), 18-33.

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance, 32*(8), 1541-1551.

Agarwal, V., & Taffler, R. J. (2007). Twenty-five years of the Taffler z-score model: Does it really have predictive ability? *Accounting and Business Research, 37*(4), 285-300.

Alifiah, M. N. (2014). Prediction of Financial Distress Companies in the Trading and Services Sector in Malaysia Using Macroeconomic Variables. *Procedia - Social and Behavioral Sciences, 129*(Supplement C), 90-98. doi:https://doi.org/10.1016/j.sbspro.2014.03.652

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

Altman, E. I. (1983). Why businesses fail. *Journal of Business Strategy, 3*(4), 15-21.

Altman, E. I. (2000). Predicting financial distress of companies: revisiting the Z-score and ZETA models. *Stern School of Business, New York University*, 9-12.

Altman, E. I., & Hotchkiss, E. (2006). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (Vol. 289): John Wiley & Sons.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting, 28*(2), 131-171.

Asgarnezhad Nouri, B., & Soltani, M. (2016). Designing a bankruptcy prediction model based on account, market and macroeconomic variables (Case study: Cyprus Stock Exchange). *Iranian Journal of Management Studies, 9*(1), 125-147.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting Review, 38*(1), 63-93.

Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance, 40*(Supplement C), 432-442. doi:https://doi.org/10.1016/j.jbankfin.2013.12.013

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

Beaver, W. H. (1968). Alternative Accounting Measures As Predictors of Failure. *Accounting Review, 43*(1), 113.

Beaver, W. H., McNichols, M. F., & Jung-Wu, R. (2005). Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy. *Review of Accounting Studies, 10*(1), 93-122. doi:10.1007/s11142-004-6341-9

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies, 21*(3), 1339-1369.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition, 30*(7), 1145-1159. doi:https://doi.org/10.1016/S0031-3203(96)00142-2

Brooks, C. (2014). *Introductory econometrics for finance*: Cambridge university press.

Carling, K., Jacobson, T., Lindé, J., & Roszbach, K. (2007). Corporate credit risk modeling and the macroeconomy. *Journal of Banking & Finance, 31*(3), 845-868. doi:https://doi.org/10.1016/j.jbankfin.2006.06.012

Casiraghi, L., & Monteiro, A. (2016, 18 April 2016). Edcon stops servicing its $950m debt in desperate bid to avoid bankruptcy. Retrieved from http://www.biznews.com/briefs/2016/04/18/edcon-stops-servicing-its-950m-debt-in-desperate-bid-to-avoid-bankruptcy/

Cassim, R. J. (2014). *The value of financial ratio analysis in predicting the failure of JSE listed companies.* (Magister Commercii), North-West University.

Castagnolo, F., & Ferro, G. (2014). Models for predicting default: towards efficient forecasts. *The Journal of Risk Finance, 15*(1), 52-70.

Charalambakis, E., & Garrett, I. (2016). On the prediction of financial distress in developed and emerging markets: Does the choice of accounting and market information matter? A comparison of UK and Indian Firms. *Review of Quantitative Finance & Accounting, 47*(1), 1-28. doi:10.1007/s11156-014-0492-y

Charitou, A., Dionysiou, D., Lambertides, N., & Trigeorgis, L. (2013). Alternative bankruptcy prediction models using option-pricing theory. *Journal of Banking & Finance, 37*(7), 2329-2341. doi:https://doi.org/10.1016/j.jbankfin.2013.01.020

Christidis, A., & Gregory, A. (2010). Some new models for financial distress prediction in the UK.

Das, S. R., Hanouna, P., & Sarin, A. (2009). Accounting-based versus market-based cross-sectional models of CDS spreads. *Journal of Banking & Finance, 33*(4), 719-730.

FTSE, R. (2017). Industry Classification Benchmark (Equity) version 2.3, 22. Retrieved from Industry Classification Benchmark (ICB) website:

Grice, J. S., & Dugan, M. T. (2003). RE-ESTIMATIONS OF THE ZMIJEWSKI AND OHLSON BANKRUPTCY PREDICTION MODELS. *Advances in Accounting, 20*(Supplement C), 77-93. doi:https://doi.org/10.1016/S0882-6110(03)20004-3

Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.): McGraw-Hill/Irwin.

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

Hlahla, B. F. (2011). *Assessing corporate financial distress in South Africa.* (Master of Management in Finance & Investment), University of the Witwatersrand.

Hol, S. (2007). The influence of the business cycle on bankruptcy probability. *International Transactions in Operational Research, 14*(1), 75-90.

Jackson, R. H., & Wood, A. (2013). The performance of insolvency prediction and credit risk models in the UK: A comparative study. *The British Accounting Review, 45*(3), 183-202.

Kidane, H. W. (2004). *Predicting financial distress in IT and services companies in South Africa.* (Magister Commercii), University of the Free State.

Kleinbaum, D. G., & Klein, M. (2010). *Logistic Regression: A Self-Learning Text* (3 ed.): Springer-Verlag New York.

Kollár, B., & Gondžárová, B. (2015). Comparison of current credit risk models. *Procedia Economics and Finance, 23*, 341-347.

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

Le, H. H., & Viviani, J.-L. (2017). Predicting bank failure: An improvement by implementing machine learning approach on classical financial ratios. *Research in International Business and Finance*.

Lin, T.-H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, probit and neural networks models. *Neurocomputing, 72*(16), 3507-3516.

Masekesa, W. M. (2011). *Macroeconomic determinants of corporate failures in South Africa (1994-2009).*

Metz, C. E. (1978). *Basic principles of ROC analysis.* Paper presented at the Seminars in nuclear medicine.

Mihalovic, M. (2016). Performance Comparison of Multiple Discriminant Analysis and Logit Models in Bankruptcy Prediction. *Economics & Sociology, 9*(4), 101.

Muller, G., Steyn-Bruwer, B., & Hamman, W. (2009). Predicting financial distress of companies listed on the JSE-a comparison of techniques. *South African Journal of Business Management, 40*(1), 21-32.

Naidoo, S. R., & Du Toit, G. (2007). A predictive model of the states of financial health in South African businesses. *Southern African Business Review, 11*(3), 33-55.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131.

Pretorius, M. (2008). Critical variables of business failure: a review and classification framework: management. *South African Journal of Economic and Management Sciences, 11*(4), 408-430.

Pretorius, M. (2009). Defining business decline, failure and turnaround: A content analysis. *The Southern African Journal of Entrepreneurship and Small Business Management, 2*(1).

Rama, K. D. (2012). *An empirical evaluation of the Altman (1968) failure prediction model on South African JSE listed companies.*

Reisz, A. S., & Perlich, C. (2007). A market-based framework for bankruptcy prediction. *Journal of Financial Stability, 3*(2), 85-131. doi:https://doi.org/10.1016/j.jfs.2007.02.001

Rose, P. S., Andrews, W. T., & Giroux, G. A. (1982). Predicting business failure: A macroeconomic perspective. *Journal of Accounting, Auditing and Finance, 6*(1), 20-31.

Shevel, A., & Rose, R. (2017, 15 June 2017). The death of Stuttafords. Retrieved from https://www.businesslive.co.za/fm/features/2017-06-15-the-death-of-stuttafords/

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business, 74*(1), 101.

Smith, M., & Liou, D.-K. (2007). Industrial sector and financial distress. *Managerial Auditing Journal, 22*(4), 376-391.

Steyn-Bruwer, B., & Hamman, W. (2006). Company failure in South Africa: classification and prediction by means of recursive partitioning. *South African Journal of Business Management, 37*(4), 7-18.

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *The Canadian Journal of Psychiatry, 52*(2), 121-128.

Sun, J., Li, H., Huang, Q.-H., & He, K.-Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems, 57*(Supplement C), 41-56. doi:https://doi.org/10.1016/j.knosys.2013.12.006

Tian, S., & Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics & Finance*. doi:https://doi.org/10.1016/j.iref.2017.07.025

Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis, 30*, 394-419.

Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *The journal of finance, 59*(2), 831-868.

Wang, J., & Dwyer, D. W. (2010). *Moody's Analytics RiskCalc v3.2 South Africa*. Modelling methodology from Moody's KMV. Analytics. Moody's.

Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics, 6*(1), 34-45. doi:https://doi.org/10.1016/j.jcae.2010.04.002