Statistical Analysis of Accounting and Financial Ratios as Determinants of Bankruptcy: A Dynamic Approach

Emmanuel Okon

A Thesis

In the Department

of

Business Analytics and Technology Management

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science (Business Analytics and Technology Management) at

Concordia University

Montréal, Québec, Canada

December 2025

© Emmanuel Okon, 2025

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Emmanuel Okon, 2025

Entitled: Statistical Analysis of Accounting and Financial Ratios as Determinants of Bankruptcy: A Dynamic Approach

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Business Analytics and Technology Management)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

____
ChairDr. Dongliang Sheng

______________________________ Examiner
Dr. Rustam Vahidov

______________________________ Supervisor
Dr. *Salim Lahmiri*

Approved by ______________________________
Dr. Suchit Ahuja, Graduate Program Director

December 2025 ______________________________
Dean of Faculty

Abstract

The current study investigates the time-varying explicatory potential of accounting data and financial ratios as explainer of bankruptcy by corporations. Using quantitative research methodology, the analysis utilizes a comprehensive panel data-set of 65,000 firm-year observations made by Polish manufacturing firms (2000-2013) to estimate Logit and Probit regression models of the 64 financial ratios, using quite rigorous variable selection procedures in order to guarantee the robustness. The results show a good degree of temporal consistency in bankruptcy explanation: there was a common set of indicators, Working Capital, Operating Profit Margin, adjusted gross margin and Retained Earnings to Total Assets, that are significant variables throughout the period. More importantly, the explanatory ability of such ratio's changes over time, with leverage indicators giving an early warning, whereas profitability and liquidity ratios dominate in the pre-bankruptcy period. These findings illustrate the need for dynamic modelling to clarify bankruptcy more accurately and offer valuable insights for developing early warning systems on a stage-by-stage basis.

Keywords: Bankruptcy, Financial Ratios, Dynamic Analysis, Logit Regression, Probit Regression.

Acknowledgement

I would like to deeply thank my supervisor, **Prof. Salim Lahmir**, who has greatly assisted and stood with me throughout the duration of this research. His critical feedback and support played a crucial role in shaping the project from the very beginning to its completion. I am highly grateful that he always made time to speak with me and consistently provided clear and well-guided direction during our conversations.

I also cannot forget my Course Adviser, **Prof. Suchit Ahuja**, who gave me my first encouragement and push as an international student arriving late and in a chaotic situation. My sincere appreciation also goes to the team of examiners, **Prof. Dongliang Sheng (Chair)** and **Prof. Rustam Vahidov**, for their valuable time and contribution.

I owe a great debt of gratitude to my family, who have been very supportive and understanding. They have always seen me as capable, and this has been my source of constant strength, enabling me to defy odds at any given moment and pursue my goals with resilience. My heartfelt appreciation goes especially to my spouse, **Mrs. Kate Clement**, whose patience and support—even from miles away—have been tremendous.

Table of Contents

Abstract iii
Acknowledgement iv
Chapter One - Introduction 1
1.1 Background of the Study 1
1.2 Problem Statement 2
1.3 Aim and objectives 3
1.3.1 Objectives: 3
1.4 Research questions 3
1.5 Contributions 3
1.6 Significance of study 4
Chapter Two - Literature Review 5
2.1 Overview 5
2.2 Theoretical Framework 5
2.2.1 Decision Usefulness Approach 5
2.3 Financial Ratios 6
2.4 Financial Distress Models 7
Statistical Models 8
2.4.1.1 Altman's Z-Score Model 8
2.4.1.2 Logistic Regression Model 8
2.4.1.3 Probit Regression Model 9
2.5 Review of Insolvency Risk Analysis Using Financial Ratios 9
2.6 Limitation of Literature 12
Chapter Three - Methodology 13
3.1 Introduction 13
3.2 Research Design 13
3.3 Data Collection 14
3.3.1 Data Source 14
3.3.2 Sampling Procedure 14
3.3.3 Variable Construction 15
3.4 Analytical Framework 17
3.4.1 Regression Models 17
3.4.2 Model Estimation 18
3.4.3 Ratio Significance 19
3.4.4 Temporal Stability Models 19

3.4.5 Model Effectiveness 19
3.4.6 Hosmer-Lemeshow (H-L) Goodness-of-Fit Test 20
Chapter Four - Data Description and Analysis 22
4.1 Descriptive Statistics and Data Characteristics 22
4.1.1 Temporal Distribution and Sample Composition 22
4.1.2 Financial Ratio Distribution Analysis 24
4.2 Variable Selection and Model Preparation 28
4.3 Logit Regression Analysis: Initial and Filtered Model Comparison 30
4.4.1 Model Performance and Fit 31
4.4.2 Key Explainers of Bankruptcy 33
Several variables consistently appeared as significant Explainers across multiple 33
4.4.3 Economic Interpretation of Coefficients 34
4.4.4 Classification Performance and Limitations 34
4.5 Probit Regression Analysis: Key Findings and Comparison with Logit Models 36
4.5.1 Model Performance and Fit 37
4.5.2 Key Explainers and Economic Interpretation 39
4.5.4 Comparison with Logit Models 40
4.5.5 Multicollinearity and Robustness 40
4.6 Limitations and Recommendations 41
4.7 Model Performance Evaluation 41
4.7.1 Goodness-of-Fit Metrics 41
4.7.2 Classification Accuracy 41
4.8 Stability and Consistency of Key Explanatory Ratios Over Time 42
4.8.1 Consistency of Significant Ratios 42
Chapter Five - Discussion of Key Findings 43
5.1 Significant Financial Ratios and Comparison with Existing Literature 43
5.2 Shifting Explanatory Value of Financial Ratio Over Time 44
5.3 Model Explanatory Performance 44
5.2 Theoretical Implications (Logit & Probit Models) 45
5.3 Research Limitations 46
5.3.1 Data-Related Constraints 46
5.3.2 Methodological Considerations 46
5.3.3 Generalizability Issues 47
5.3.4 Implementation Challenges 47
5.4 Practical Recommendations 47
Chapter Six - Conclusion and Recommendation 49
Introduction 49

6.1 Purpose of the Study 49
6.2 Models Used and Key Contributions 49
6.3 Major Findings 50
6.3.1 Key Financial Variables 50
6.4 Recommendations 51
6.5 Future Research Directions 52
6.6 Contribution to Knowledge 52
6.7 Conclusion 53
References 54

List of Tables

Table 2.1- Summary of Empirical Studies on Financial Distress Explanation........................11
Table 3.1- List of Financial Ratios 15
Table 4.1- Year 1 statistical 25
Table 4.2 - Year two statistical 25
Table 4.3 - Year three statistical 26
Table 4.4 - Year four statistical...27
Table 4.5 - Year 5 statistical 27
Table 4.6 - Logit model comparison...30
Table 4.7 - All significant explainer 30
Table 4.8 - Confusion Matrices for Model explanations Across Years 1 to 5.........................35
Table 4.9 - Probit model comparison table 36
Table 4.10 - All significant explainer 36

List of Figures

Figure 4.1 - Year 1 bankruptcy distribution...22
Figure 4.2 - Year 2 bankruptcy distribution...22
Figure 4.3 - Year 3 bankruptcy distribution...23
Figure 4.4 - Year 4 bankruptcy distribution...23
Figure 4.5 - Year 5 bankruptcy distribution...24
Figure 4.6 - Heatwave Analysis of Logit P-Values 32
Figure 4.7 - Heatwave Analysis of Probit P-Values 38

Chapter One - Introduction

1.1 Background of the Study

The end of corporate existence becomes inevitable when a firm reaches bankruptcy after enduring continued financial trials which put its survival at risk. Bankruptcy affects both the organization directly and numerous stakeholder groups, which include creditors and investors as well as employees, suppliers, and regulators (Bernstein et al., 2017). Companies face bankruptcy, which leads creditors to suffer uncollected debts while investors lose their capital and employees lose their jobs. At the same time, the economy experiences lower production along with weakened financial systems (Kolay et al., 2016).

Historically, accounting and financial ratios remain a fundamental evaluation and financial health assessment tool in business practice (Kadim, Sunardi, and Husain, 2020). The empirical research of Beaver (1966) showed that financial ratios possessed the capability to become useful indicators for organisational distress explanation. The models developed during that period were characterised by their static structure and relied on identifying common attributes associated with financially troubled firms.

According to Andirawan and Salean (2016), ratio analysis serves as a practical method for determining whether financial data can be used to classify or predict bankruptcy. To reduce or avoid financial distress, companies can examine the financial status by balance sheets and income statements by the means of the financial analysis techniques (Pernamasari, 2020). Borrowed from these statements, the ratios provide some meaning about a firm's liquidity, profitability, leverage and its operational efficiency, major dimensions of financial stability (Rashid, 2020).

A major breakthrough at the time presented itself in 1968 when Edward Altman developed the Z-score model that used multivariate discriminant analytics to integrate a set of financial ratios into a single measure of predictability (Jarvis, 2024). Altman's approach showed that unhealthy firms in danger of bankruptcy could be properly differentiated from normal (healthy) firms through a rational combination of financial indicators (Asif et al., 2024).

1.2 Problem Statement

Despite years of research, a precise measurement of the bankruptcy risk is still complex, especially in changing economic and industry conditions (Ohlson, 1980). Financial and accounting ratios, including ones that measure liquidity, profitability, leverage and efficiency, have traditionally been important in bankruptcy risk assessment because of their power to condense complex financial data into actionable intelligence (Beaver, 1966; Yazdanfar and Nilsson, 2008). These ratios offer a picture of how financially healthy a firm is at levels of very early warning signals, for example, the decrease of cash flows or the increase of the debt load, which allows for timely intervention (Hillegeist et al., 2004).

The significance of financial ratios is based on the ability to identify critical aspects of firm operations in numerical form, providing quantifiable indicators of short-term solvency and long-term viability (Altman, 1968). For example, liquidity ratios, such as the current ratio, indicate ability of a firm to honour its immediate obligations, while leverage ratios, such as debt-to-equity, illustrate the dangers of excessive borrowing, which is often a forerunner of bankruptcy (Sun et al. 2016). In the past, integrating more than one ratio has been demonstrated to improve the analysis of bankruptcy risk as compared to prior single ration methods (Xu et al., 2014).

However, there are significantly different choices of financial ratios as well as the number of ratios presented in the studies because there is no general benchmark on the number of financial ratios to use (Zaini and Mahmuddin, 2019). For instance, Bose (2006) used 24 ratios, Wang and Chen (2006) used 11, Hua et al. (2007) included 22, Zaini and Mahmuddin (2019) employed 24, and Habibi and Iqbal (2021) used as many as 47. In many cases, models incorporate large sets of ratios, which, while comprehensive, can obscure key insights and reduce the overall effectiveness of bankruptcy explanation (Dimitras, Zanakis and Zopounidis, 1996; Xu et al., 2014). This over-reliance on broad sets of ratios risks diluting the focus on critical indicators, making it harder to identify the most relevant signs of financial distress. This gap points to the need for a refined approach to leverage financial ratios effectively in bankruptcy risk assessment, focusing on the most impactful ratios that offer the clearest indicators of financial health and potential distress.

1.3 Aim and objectives

Hence, this study aims to determine key financial indicators that explain corporate bankruptcy risk through dynamic regression analysis across different economic conditions.

1.3.1 Objectives:

To achieve this aim, the study is structured around the following objectives:

1. Identify and analyze the accounting and financial ratios that demonstrate the risk of bankruptcy.
2. Assess whether the accounting and financial variable of bankruptcy remain stable over time.
3. Evaluate the explanatory power of accounting and financial ratios in explaining of bankruptcy.

1.4 Research questions

In line with the aim and objectives of the study, the following research questions are posed to further guide the investigation:

1. Which accounting and financial ratios significantly explain bankruptcy risk?
2. Are the accounting and financial variables of bankruptcy risk stable over time?
3. How well do significant accounting and financial determinants explain the likelihood of bankruptcy?

1.5 Contributions

The research contributes to the body of knowledge on financial distress explanation by proposing a dynamic model that examines the effect of financial ratios on bankruptcy risk over five (5) years in direct response to the temporal shortcoming of the static models, which are widespread in the literature. The method of monitoring the stability of financial variables across economic cycles allows for increasing the accuracy of bankruptcy explanations in fluctuating real-life conditions.

To guarantee methodological soundness, we complementarily contrast Logistic regression assuming logistic distribution and Probabilistic regression assuming normal distribution models and offer essential information on what distribution assumptions can have on the explanatory power of financial variables and lay the groundwork of further model selection. These analytical developments are powered by an immensely detailed database of 65,000 firm-year observations, including more than 10,000 well-performing firms and more than 700 bankruptcies with multi-year pre-filing data, a magnitude that greatly enhances statistical capacity and generalizability compared to the earlier studies with small sample sizes. Most importantly, the results revealed some key financial ratios that had a robust consistent effect on the bankruptcy risk, providing practical recommendations to risk managers and bankers to focus on high-signal indicators and perfect early-warning systems.

1.6 Significance of study

The insight from this research aims to support current and future firms in proactively managing financial risk, as it offers practical relevance to corporations by highlighting specific ratios that can aid in monitoring financial health and improving decision-making under certain economic conditions. Furthermore, the findings have wider implications for corporate governance and financial regulation as they provide valuable input for policymakers and regulators to enhance financial reporting standards and design measures that can help mitigate the broader economic consequences of bankruptcy, including job losses, supply chain disruptions and systemic stability.

Chapter Two - Literature Review

2.1 Overview

This chapter is anchored by providing its theoretical framework. The most fundamental is the Decision Usefulness Approach that helps in financial reporting that allows stakeholders to make well-informed decisions. Financial ratios are presented as viable tools to simplify and interpret a difficult financial information. The review then examines the significant advances in the explanation of bankruptcies, beginning with the initial applications of the single ratio, continuing to more sophisticated statistical models such as the Z-score of Altman, logistic regression of Ohlson and probit model of Zmijewski, which combine a number of indicators to increase explanatory accuracy. The last section provides an overview of empirical research on financial ratio use in determining bankruptcy. These arguments describe a generalized use of ratio-based models within industries throughout the years. Thus, the chapter outlines the main gaps in the literature, whereby dynamic, context sensitive models are required which change over time.

2.2 Theoretical Framework

2.2.1 Decision Usefulness Approach

Based upon the Decision Usefulness Approach, the primary goal of accounting theory is the provision of financial reporting that enables the stakeholders to make informed economic decisions. The objective of this approach, as stated by the FASB (2010), is to provide the stakeholders with relevant, reliable, understandable and comparable (Investors, creditors, employees, customers, Government agencies and the public) information. According to Scott (2009), the practical goal of this approach can be to provide the most informative reports to decision-makers, although making flawless financial statements is impossible (Jabbar, 2017).

To elaborate more, financial ratios are also instrumental to the Decision Usefulness Approach, as they transform disorganized financial information into neat, actionable information. Stakeholders obtain accurate perspectives of the solvency of a firm through the liquidity, leverage, profitability, and efficiency ratios (Henderson et al., 2006). Ratios should work by predicting future, should be reliable, should be derived by verifiable sources; should be reported clearly by firms and time (Ohlson, 1980). Ratios bind the raw numbers to what is needed by decision-makers, simplifying complex data into useful metrics, which will enable us to make an evaluation of options in cases where there is a lot of uncertainty. This motivation of smart decision-making under uncertainty is directly related to the decision theory, which provides an additional layer to the Decision Usefulness Approach demonstrating the way rational decisions are made. The single-person decision theory by Stinchcombe (2000) demonstrates the way personal beliefs are honed by individuals through new knowledge and considering the probable consequences (Scott, 2009). Healthy ratios signal future stability, pushing investors to greater risk-free returns, whereas unhealthy ratios caution us to be cautious and minimize risk (Salome et al., 2021).

2.3 Financial Ratios

Financial ratios are a quantifiable association of the facts in the financial statements of an organization that yields the feasibility of its operations and finances. They are important measurements of performance and determination of strengths and weaknesses and any possible risks. Ratios are used to offer comparisons within a company over a period of time or with a company in related sectors through the comparison of data in financial statements and are therefore fundamental to economic analysis (Lan, 2012). To illustrate, Altman (1968) makes use of ratios to predict insolvency in business, whilst some other researchers have used it to determine future profitability. Ratio analysis typically applies to two methods such as the ratio of a company to industry rates or peers, and the changes in the ratios of a company over a number of periods (Miswanto, Kusumasari, and Anggoro, 2018).

The idea of financial ratios was invented in the late 19th century when the current ratio was created, and then such inventions as the DuPont model, which breaks down the return on equity to demonstrate how the companies create shareholder value, were invented (Mankin and Jewell, 2014). Ratios need a reference point to obtain their maximum utilities, which may be the industry averages or the essence of the business of the company.

The adaptability of financial ratios has led to an expanding array of metrics tailored to specific analytical needs. This however, suffers from proprietary shading variants that result in unpredictable naming conventions and overlapping formulas, which confuses users. The convergence of complex financials necessitated by complex accounting standards and fixed data presentation contributes to further complexity in ratio analysis. Ratios are generally categorized into univariate and multivariate analyses (Kusumasari, Anggoro & Miswanto, 2018). While univariate analysis attends to a single measure to describe the state of a company, multivariate analysis combines the measures from four key categories: liquidity, solvency, efficiency, and profitability.

Liquidity ratios evaluate the ability of a company to pay its near-term liabilities, including current, quick, cash and inventory-to-working-capital ratios (Horne & Wachowicz, 2009). Increasing liquidity ratios indicate greater short-term financial stability, the current ratio shows the capability to liquidate short-term debt from current assets. Solvency ratios assess how much a firm depends on debts to finance its operations and consist of debt-to-assets ratio, debt-to-equity ratio and long-term debt-to-equity ratio (Ross, 2016). An upswing in these ratios specifically indicates growing debts, and invariably it is due to a destabilized state of finances.

Efficiency ratios, also referred as activity ratios, measure how efficiently a company is using these resources, i.e cash turnover, accounts receivable turnover, inventory turnover and total asset turnover (Titman, Keown, and Martin 2014). Efficiency ratios at higher levels imply improved utilization of resources providing useful information to stakeholders. There are profitability ratios which measure the companies' ability to earn from their strategies and capital such as earnings per share (EPS), the net profit margin (NPM) the return on investment (ROI) and return on equity (ROE) (Brealey, Myers and Marcus 2012).

2.4 Financial Distress Models

Bankruptcy forcasting is commonly associated with the science of explaining financial failure (Meeampol et al., 2016) and will always remain a top study in finance and accounting. The importance of such an analysis is found in the fact that it allows informing such stakeholders as the investors, creditors, and regulators what is the financial stability of the company and the possible risk of default. These models are very critical in determining the probability of financial distress among public firms (Meeampol et al., 2016).

Before the advent of quantitative techniques, creditworthiness was primarily assessed through qualitative reports issued by specialized agencies. These evaluations typically summarize a firm's historical performance to aid decision-making (Altman, 1968). In the 1930s, early efforts in financial distress evaluation employed univariate discriminant analysis, where analysts relied on a single financial ratio, such as the current ratio, to assess a firm's financial position (Altman, 1968; Beaver, 1966). Notable studies by Fitzpatrick (1931) and Merwin (1942) compared the financial data of bankrupt and solvent firms in an attempt to identify early warning indicators. However, a single financial ratio was insufficient, as most often it did not provide a complete financial overview (Bellovary et al., 2007).

The negative sides of the univariate analysis became apparent because simple financial measures were often leading to incorrect conclusions. While a poor profitability measure might mean possible instability, a firm's liquidity can sometimes cushion against this (Altman, 1968). An extensive univariate analysis (Beaver, 1966) identified that ratios pertaining to cash flow greatly increase the ability to identify signals of distress. Nevertheless, relying on a single indicator constrained the accuracy of explanations, highlighting the need for a more robust analytical framework.

Statistical Models

2.4.1.1 Altman's Z-Score Model

To advance analytical research, Altman (1968) created the multivariate discriminant analysis through his work on the Z-score model. The unified explanatory formula of the Z-score model combined financial ratios which measured liquidity along with profitability and leverage. Through incorporating various indicators, the Z-score model achieved better reliability for short-term bankruptcy forecasts while establishing financial health ranges from healthy to distressed and intermediate levels (Heine, 2000).

In Altman's model, the z-value represents a company's financial score used to assess the likelihood of bankruptcy. A higher z-score suggests financial stability, while a lower score indicates potential bankruptcy. Specifically, companies with a z-score below 1.81 are classified as bankrupt, while those with a score above 2.99 are considered financially healthy. Scores falling between 1.81 and 2.99 fall into a gray area, where the projection is less certain. A score of 2.675, for example, implies an equal probability (50%) of bankruptcy or survival. (Ezejiofor and Okerekeoti, 2021)

Many other researchers have developed general models based on multivariate discriminant analysis that can be used by any enterprise, such as Bilderbeek (1979), Blum (1974), Daniel (1968), Deakin (1972), Laitinen (1991), Lussier et al. (1996), and many others. However, because these models cannot predict how likely a firm is to fail, their explanatory power is partially limited. Similarly, most studies in the past have shown that the basic assumptions for the multiple discriminant analysis are often violated. Considering these problems, newer methods involving logit and probit analysis were created over time, which do not require these assumptions to be met (Gajdosikova and Valaskova, 2023).

2.4.1.2 Logistic Regression Model

Ohlson (1980) introduced conditional logistic regression as an alternative to multivariate discriminant analysis for bankruptcy prediction. Unlike earlier models, Ohlson's approach focused on identifying which financial ratios significantly influence bankruptcy risk using a substantially larger dataset of 2,000 firms, thereby improving the model's generalizability. The model outputs a probability, referred to as the o-score, which estimates the likelihood of bankruptcy on a scale from 0 to 1. Companies with scores closer to 1 are considered more likely to go bankrupt.

Logistic regression integrates multiple firm characteristics into a single probabilistic estimate, making it more flexible in handling various types of financial information. Critics point out that logistic models are sensitive to multicollinearity, depend heavily on the assumption of a logistic distribution, and can be affected by missing values and outliers (Balcaen and Ooghe, 2006). Additionally, the binary classification approach, where firms are labeled as either bankrupt or not, may oversimplify financial distress, especially when initial conditions already suggest insolvency (Abdullah, Ahmad, and Rus, 2008).

2.4.1.3 Probit Regression Model

Similar to the logistic regression, the probit regression handles binary values because the person goes bankrupt or not but the cumulative standard normal distribution is used in its place of the logistic func-tion (Zmijewski, 1984). According to this model, there is a different meaning to its probability score such that companies with a probability score lower than 0.5 remain financially stable, but a firm with a probability score of 0.5 or higher is vulnerable to financial risks. Such external factors as the firm size, in combination with industry, have significant roles in the estimation of the risk of bankruptcy based on the model of Zmijewski (Zubanovic and Ahmeti, 2020).

The simplicity of the interpretation of both the logit and probit model can be listed among the advantages: the models provide a concise estimation of the likelihood of a firm going bankrupt in the future (Gajdosikova and Valaskova, 2023). Since they are popular and easy to interpret, many studies have tried to compare their performance with previous explanatory methods (Gajdosikova and Valaskova, 2023). As an example, Lennox (1999) compared models constructed using multiple discriminant analysis (MDA) with the one constructed using conditional probability methods.

2.5 Review of Insolvency Risk Analysis Using Financial Ratios

Bhunia and Sarkar (2011) applied financial ratios and multiple discriminant analysis to predict distress in Indian pharmaceutical firms, highlighting the significance of liquidity and profitability ratios. Yap et al. (2011) used a univariate approach in Malaysia, noting significant differences in financial ratios between failed and non-failed firms, with the cash flow-to-total debt ratio being a key indicator of solvency. Maricica and Georgeta (2012) identified significant differences in profitability, financial position, and leverage ratios between performing and non-performing firms. Hassani and Parsadmehr (2012) found debt-to-equity, net profit-to-net sales, and working capital-to-assets ratios significant in predicting financial crises.

Yap et al. (2012) identified cash flow-to-total debts, cash-to-current liabilities, total debts-to-total assets, and retained earnings-to-total assets ratios as significant determinant of business failure. Mondal and Roy (2013) noted that the profit-after-tax growth rate and debt-to-equity ratio were critical for identifying sick companies. Keener (2013) highlighted lower cash-to-current liability ratios, reduced cash flow margins, and higher debt-to-equity ratios as significant in firm failure. Alifiah (2014) found that debt ratio, total assets turnover, working capital, and net income-to-total assets ratios were significant for Malaysian firms.

AlKassar and Soileau (2014) applied Altman's Z-score model, identifying seven significant ratios. Mraihi (2015) developed a logistic regression model with twelve ratios, emphasizing liquidity and solvency over profitability ratios. Zohra et al. (2015) used three ratios, net working capital-to-owner's equity, accounts receivable turnover, and owner's equity-to-fixed assets, to differentiate distressed firms. Jabeur (2017) validated partial least squares logistic regression for distress explanation. Agrawal and Maheshwari (2019) found industry beta significant in identifying defaults among Indian firms. Balasubramanian et al. (2019) identified net asset value, long-term debt-equity ratio, and other financial and non-financial factors as key distress pointers. Ogachi et al. (2020) concluded that inventory turnover, asset turnover, debt-equity, debtors' turnover, current, and working capital ratios were critical for bankruptcy.

Many studies have shown the usefulness of financial ratios in projecting bankruptcy (Altman, 1968; Ohlson, 1980; Bhunia and Sarkar, 2011; Alifiah, 2014), but the selection of ratios often relies on prior research rather than a comprehensive evaluation of their explanatory reliability (Karels and Prakash, 1987). The choice of variables, both which and how many, is important (Balcaen and Ooghe, 2006), as it could further simplify the analysis. Despite the wide range of available variables, there is no consensus on which financial ratios best distinguish between bankrupt and non-bankrupt companies, as many statistical models are static and fail to evaluate the stability of these ratios over time (Mraihi, 2015; Jabeur, 2017). This study addresses these gaps by adopting a dynamic approach to identify the most impactful financial ratios and assess their consistency across time.

Table 2.1- Summary of Empirical Studies on Financial Distress Explanation

Study	Data / Sample	Input Variables	Models Used	Main Finding
Bhunia and Sarkar (2011)	Pharmaceutical firms in India	Financial ratios	Multiple discriminant analysis	Liquidity and profitability ratios were significant Explainers of financial distress.
Yap et al., (2011)	64 listed companies in Malaysia over 10 years	Financial ratios	Univariate statistical analysis	Cash flow to total debt ratio effectively predicted business failure.
Maricica and Georgeta (2012)	Romanian companies	Profitability, financial position, leverage ratios	Comparative statistical analysis	Identified differences in financial indicators between healthy and distressed firms.
Hassani and Parsadmehr (2012)	Iranian firms	Debt to equity ratio, net profit to net sales, working capital to total assets	Financial ratio analysis	Found strong indicators of financial crisis among selected ratios.
Yap et al., (2012)	Malaysian firms	Cash flow and solvency ratios	Financial ratio analysis	Key financial ratios were effective in business failure explanation.
Mondal and Roy (2013)	Indian companies	Financial ratios	Comparative financial analysis	Clearly distinguished between distressed and healthy companies.
Keener (2013)	United States public companies	Liquidity and leverage ratios	Financial ratio analysis	Low liquidity and high leverage were linked to greater financial risk.
Alifiah (2014)	Trading and services sector companies in Malaysia	Debt ratio, total assets turnover, net income to total assets	Logistic regression model	Selected financial ratios significantly explained the likelihood of failure.
AlKassar and Soileau (2014)		Seven financial ratios across categories	Altman's Z-score model	Confirmed the model's reliability in financial distress explanation.
Mraihi (2015)		Liquidity, solvency, and profitability ratios	Logistic regression model	Liquidity and solvency were stronger Explainers than profitability.

Zohra et al. (2015)	Pakistani firms	Working capital ratio, equity ratio, fixed asset ratio	Financial ratio comparison	Demonstrated differences in financial structure between firm types.
Jabeur (2017)		Financial ratios	Partial least squares logistic regression	Confirmed model accuracy in explaining financial distress.
Agrawal and Maheshwari (2019)	Indian corporate sample	Industry-specific risk factors (e.g., beta)	Econometric modeling	Industry-related risk was a significant Explainer of corporate default.
Balasubramanian et al. (2019)	Indian firms	Net asset value, debt-to-equity ratio, and non-financial variables	Mixed methodological approach	Non-financial information improved Explanatory reliability.
Ogachi et al. (2020)		Turnover ratios and liquidity measures	Financial ratio analysis	Specific financial indicators were effective in distress identification.

2.6 Limitation of Literature

While the literature review offered an overview of existing research, several areas remain underexplored in the current study.

The study excluded non-financial variables such as management quality, market conditions, or corporate governance, and it did not consider industry-specific models. Also, the review did not cover advanced statistical methods and machine learning techniques, such as neural networks or decision trees, which have gained popularity in bankruptcy explanation. Furthermore, the study focused only on traditional statistical models, such as Altman's Z-score and logistic regression. Additionally, Country-specific contexts and regulatory environments were not included in the research. Furthermore, the study did not explore the impact of corporate restructuring, government interventions, or other external factors that may alter the bankruptcy explanation process.

Chapter Three - Methodology

3.1 Introduction

The methodological framework is delineated to investigate the degree of explanatory power of financial ratios in the evaluation of bankruptcy risk. Using the quantitative technique of analysis, the study examines 64 financial ratios through a logit and probit statistical model. The research examines data from Emerging Markets Information Service (EMIS) issued between 2000 and 2013 to review bankrupt and non-bankrupt firms. The research method systematically looks at the stability as well as the determining ability of these ratios over time

3.2 Research Design

The research design adopted in the study is a quantitative explanatory research design that investigates the explanatory capabilities of financial ratios towards determining bankruptcy risk. This design is best for the analysis of large-scale financial data, scientific modelling of correlation between the independent variables (financial ratios) and a dichotomous dependent variable (bankruptcy status) (Babbie, 2013; Creswell, 2018). The study utilizes secondary data, it is easier to obtain secondary data because there is more resource eligibility as compared to primary data; it is less expensive (Patil, 2022).

Administrative datasets, in particular, commonly include the use of large sample sizes because they are systematically collected. Additionally, many sources of secondary data such as, administrative records, are accumulated during a long period of time, enabling the researchers to study longitudinal trends (Patil, 2022). The study obtains its data from standardized annual reports and audited financial statements found within a financial database to ensure that the data it comes up with is reliable and consistent.

Logistic and probit regressions are the fundamental analytical tools selected based on their capability to estimate non-linear relationship in binary data (Jose et al., 2020). These models are heavily used in bankruptcy explanation because of their interpretability and robustness. Moreover, the study uses a panel data structure integrating cross-sectional and longitudinal dimensions. This method increases the power of the models in explaining firm-specific heterogeneity and temporal dynamics, thus providing a better understanding of bankruptcy risks (Baltagi, 2005).

3.3 Data Collection

3.3.1 Data Source

Financial data for this study is obtained from the Emerging Markets Information Service (EMIS) database which is a widely use financial repository for emerging markets as is the case with Poland (EMIS, 2024). Researchers can calculate financial ratios for liquidity, profitability, leverage, and turnover, as all databases provide full balance sheet, income sheet, and cash flow statements. Availability of data on bankruptcy accessed from the database enables the researchers to find insolvent companies more precisely offering the trustworthy outcome classification.

Strength of the platform is its specialization in emerging markets and for Polish businesses, especially relevant data. Although the international databases may not provide adequate detail for minor markets, EMIS makes sure that its standardized data is tailored to the local economic setting, thereby ensuring reliability and consistency (EMIS, 2024). With comprehensive protocols that involve cross checking national registry, the database is very effective in controlling reporting inaccuracies. The availability of the database to researchers and its use in such works as Zięba et al. (2016), proves a confirmatory suitability for our work.

3.3.2 Sampling Procedure

The sample is made up of Polish manufacturing firms, between 2000 and 2013, chosen to represent the sector's high bankruptcy incidence and economic value. The emphasis on Polish manufacturing firms is based on the economic relevance of the sector and the high bankruptcy rates, creating a dynamic setup to examine ratio performance during different economic cycles, from boom in the early 2000s to the 2007–2009 financial crisis to recovery there from (Drozdowicz-Bieć, 2012; Bogdan et al. (2015; Zięba et al. (2016). 2,400 statements were generated approximately 700 companies filing for bankruptcy between 2007-2013, at least one financial statement in five (5) years before bankruptcy. This window captures a financial distress signal before bankruptcy. With more than 10,000 companies operating without bankruptcies filings from 2000 through to 2012 with each having at least three consecutive financial statements, producing an overall number of 65,000 statements (Zięba et al, 2016).

Sampling prioritized data availability and quality. Bankrupt firms were identified through EMIS bankruptcy records, verified using filing dates and legal status. Nonbankrupt firms were excluded if they declared bankruptcy, ensuring a clear distinction. Firms with incomplete data, fewer than three consecutive statements for non-bankrupt firms or no statements within five years for bankrupt firms were excluded. All firms were drawn from the manufacturing sector to mitigate industry-specific biases, and statements were standardized using EMIS's uniform templates.

3.3.3 Variable Construction

A total of 64 financial ratios from Zięba et al. (2016) serve as the evaluation framework for firm financial health to determine bankruptcy risk. The set of financial ratios was initially created for Polish manufacturing businesses to analyze liquidity alongside profitability and leverage, and turnover metrics for complete firm stability assessment (Altman, 1968; Beaver, 1966; Zięba et al., 2016). Examples include X3 (Working Capital / Total Assets) and X4 (Current Assets / Short-term Liabilities) 3 for liquidity, X1 (Net Profit / Total Assets) and X18 (Gross Profit / Total Assets) for profitability, X2 (Total Liabilities / Total Assets) and X8 (Book Value of Equity / Total Liabilities) for leverage, and X9 (Sales / Total Assets) and X60 (Sales / Inventory) for turnover. The complete list of financial ratios is presented in Table 1 below.

Ratios were calculated using standardized formulas from EMIS financial statements. For instance, X5 ([(Cash + Short-term Securities + Receivables - Short-term Liabilities) / (Operating Expenses - Depreciation)] * 365) required precise cash flow and expense data, which EMIS reliably provided. The dataset was structured as panel data, enabling both cross-sectional and longitudinal analyses to capture firm-specific and temporal variations (Sheikh and Yahya, 2015).

Table 3.1- List of Financial Ratios

ID	Description	ID	Description
X1	Net profit / Total assets	X33	Operating expenses / Short-term liabilities
X2	Total liabilities / Total assets	X34	Operating expenses / Total liabilities
X3	Working capital / Total assets	X35	Profit on sales / Total assets

X4	Current assets / Short-term liabilities	X36	(Current assets - Inventories) / Sales
X5	[(Cash + ST securities + Receivables - ST liabilities) / (Op. expenses - Depreciation)] * 365	X37	(Current assets - Inventories) / Long-term liabilities
X6	Retained earnings / Total assets	X38	Constant capital / Total assets
X7	EBIT / Total assets	X39	Profit on sales / Sales
X8	Book value of equity / Total liabilities	X40	(Current assets - Inventory - Receivables) / Short-term liabilities
X9	Sales / Total assets	X41	Total liabilities / [(Operating profit + Depreciation) * (12/365)]
X10	Equity / Total assets	X42	Profit on operating activities / Sales
X11	(GP + Extraordinary items + Financial expenses) / Total assets	X43	Rotation receivables + Inventory turnover in days
X12	Gross profit / Short-term liabilities	X44	(Receivables * 365) / Sales
X13	(Gross profit + Depreciation) / Sales	X45	Net profit / Inventory
X14	Total liabilities * 365 / Sales	X46	(Current assets - Inventory) / Short-term liabilities
X15	(Total liabilities * 365) / (Gross profit + Depreciation)	X47	(Inventory * 365) / Cost of products sold
X16	(Gross profit + Depreciation) / Total liabilities	X48	EBITDA / Total assets
X17	Total assets / Total liabilities	X49	EBITDA / Sales
X18	Gross profit / Total assets	X50	Current assets / Total liabilities
X19	Current liabilities / Total assets	X51	Short-term liabilities / Equity
X20	(Inventory * 365) / Sales	X52	(Short-term liabilities * 365) / Cost of products sold
X21	Sales (n) / Sales (n-1)	X53	Equity / Fixed assets
X22	Profit on operating activities / Total assets	X54	Constant capital / Fixed assets

X23	Net profit / Sales	X55	Working capital
X24	Gross profit (in 3 years) / Total assets	X56	(Sales - Cost of products sold) / Sales
X25	(Equity - Share capital) / Total assets	X57	(Current assets - Inventory - ST liabilities) / (Sales - GP - Depreciation)
X26	(Net profit + Depreciation) / Total liabilities	X58	Total debt / Total assets
X27	Profit on operating activities / Financial expenses	X59	Long-term liabilities / Equity
X28	Working capital / Fixed assets	X60	Sales / Inventory
X29	Logarithm of total assets	X61	Sales / Receivables
X30	(Total liabilities - Cash) / Sales	X62	(Short-term liabilities * 365) / Sales
X31	Gross profit / Sales	X63	Sales / Short-term liabilities
X32	(Current liabilities * 365) / Cost of products sold	X64	Sales / Fixed assets

3.4 Analytical Framework

3.4.1 Regression Models

As a part of Generalized Linear Models (GLM) category, logit and probit models are popular in both academic and business settings for explaining categorical responses. These models use unique link functions, including the logit function in logistic regression and the probit function in probit regression, to relate the dependent variable to a set of independent variables (Hosmer Jr. et al., 2013). Logistic and probit regression models are used to project binary bankruptcy results (1 = bankrupt, 0 = non-bankrupt), for their ability to adjust for and have low sensitivity to outliers in handling dichotomous variables (Jose et al., 2020). The logistic regression model gives an estimation of the probability of bankruptcy as:

$$P(Y_i = 1 \vee X_i) = \frac{1}{1 + exp\left(-(\beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik})\right)} = \sigma(\beta_0 + \beta^T X_i)$$

Y_i is the bankruptcy status, $X_i = (X_{i1},...,X_{ik})$ is the vector of k ratios, $\beta = (\beta_1,...,\beta_k)$ are coefficients, β_0 is the intercept, and $\sigma(z) = \frac{1}{1+exp(-z)}$ The sigmoid function.

The log-odds form is:

$$ln\left(\frac{P(Y_i = 1 \vee X_i)}{1 - P(Y_i = 1 \vee X_i)}\right) = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik}$$

Probit regression uses the standard normal Cumulative Density Function (CDF):

$$P(Y_i = 1 \vee X_i) = \Phi(\beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik}) = \Phi(\beta_0 + \beta^T X_i)$$

where Φ is the standard normal CDF. Logistic regression assumes logistic distribution of errors, and probit assumes normal distribution, with robustness of different structures of errors. Both models are compared to evaluate explanatory performance against interpretability and applicability to imbalanced data (through weighted likelihood).

3.4.2 Model Estimation

Parameters are estimated using Maximum Likelihood Estimation (MLE), maximizing the log-likelihood for logistic regression:

$$l(\beta_0, \beta) = \sum_{i=1}^{n} \left[Y_i log\left(\sigma(\beta_0 + \beta^T X_i)\right) + (1 - Y_i) log\left(1 - \sigma(\beta_0 + \beta^T X_i)\right)\right]$$

For Probit Regression:

$$l(\beta_0, \beta) = \sum_{i=1}^{n} \left[Y_i log\left(\Phi(\beta_0 + \beta^T X_i)\right) + (1 - Y_i) log\left(1 - \Phi(\beta_0 + \beta^T X_i)\right)\right]$$

To address the imbalanced dataset (700 bankrupt vs. >10,000 non-bankrupt firms), a weighted log-likelihood is used (Czepiel, 2002):

$$l_w(\beta_0, \beta) = \sum_{i=1}^{n} w_i\left[Y_i log\left(P(Y_i = 1 \vee X_i)\right) + (1 - Y_i) log\left(1 - P(Y_i = 1 \vee X_i)\right)\right]$$

Where $w_i = \frac{N}{N_+}$ for bankrupt firms (N_+ is bankrupt firms, N is total firms) and $w_i = 1$ for non-bankrupt firms. This enhances sensitivity to bankruptcies. MLE uses iterative optimisation, with robust standard errors to correct for heteroskedasticity, critical for diverse ratios.

3.4.3 Ratio Significance

Full-sample models identify significant ratios using Wald tests ($H_0 : \beta_j = 0$), with $p < 0.05$ indicating significance. Reliable inference is provided by robust standard errors. Backwards elimination ($p < 0.05$) simplifies the model, thus minimizes overfitting (Basu et al., 2016). For example, a negative coefficient of X1 (Net Profit / Total Assets), implies that higher profitability lowers bankruptcy risk

3.4.4 Temporal Stability Models

To evaluate the stability of financial ratios across economic cycles, coefficient stability is assessed using time-specific models. Multicollinearity is checked via the Variance Inflation Factor (VIF):

$$VIF_j = \frac{1}{1 - R_j^2}$$

Where $R^2{}_j$ is the R-squared from regressing ratio j on all other ratios. Ratios with VIF > 10 are removed to ensure model reliability (Murray et al., 2012).

3.4.5 Model Effectiveness

Model performance is evaluated using multiple metrics. McFadden's R^2 is calculated as (Hu et al., 2006):

$$R^2_{McFadden} = 1 - \frac{l(\beta_0, \beta)}{l_0}$$

where l_0 is the log-likelihood of the intercept-only model. Values between 0.2 and 0.4 indicate a good fit. The Likelihood Ratio Test is computed as:

$$LR = -2\left(l_0 - l(\beta_0, \beta)\right)$$

with significant LR values confirming model improvement over the null model. AIC and BIC are given by:

$$AIC = -2l(\beta_0, \beta) + 2k, BIC = -2l(\beta_0, \beta) + kln(n)$$

where k is the number of parameters and n is the sample size. Lower values indicate better model fit. Classification metrics include sensitivity (true positive rate), specificity (true negative rate), and Area Under the Curve (AUC). An AUC > 0.8 suggests strong discriminatory power, critical for imbalanced datasets (Jose et al., 2020).

3.4.6 Hosmer-Lemeshow (H-L) Goodness-of-Fit Test

Hosmer-Lemeshow (H-L) test is usually applied to assess the calibration of the logistic regression model used for explaining bankruptcy risk. Calibration means alignment of the predicted probabilities with the observed binary outputs. H-L test is used to estimate the goodness of fit in cases of binary classification models because it compares the observed and expected rates of division of the population into groups (Hosmer and Lemeshow, 1980). For every subgroup, the test calculates the mean predicted probability and then multiplies it by the number of firms, and finds the number of expected bankruptcies and non-bankruptcies and compares these with observed frequencies. The H-L statistic then computed is (Fagerland and Hosmer, 2012; Surjanovic, Lockhart and Loughin, 2020):

$$\chi^2_{HL} = \sum_{j=1}^{g} \left(\frac{(O_{1j} - E_{1j})^2}{E_{1j}} + \frac{(O_{0j} - E_{0j})^2}{E_{0j}}\right)$$

Let g den the number of groups. The observed counts of bankrupt and non-bankrupt firms in group j are represented by $O_{1j}$and O_{0j} respectively. The corresponding expected counts, based on the model's predicted probabilities, are denoted by $E_{1j}$and E_{0j}. The expected frequencies are calculated using the mean predicted probability $\overline{p}_j$ in each group, such that:

$$E_{1j} = n_j \cdot \overline{p}_j \text{and } E_{0j} = n_j \cdot (1 - \overline{p}_j),$$

where n_j is the number of firms in group j. Under the null hypothesis, the Hosmer-Lemeshow statistic approximately follows a chi-square distribution with g−2 degrees of freedom.

The output of test contains Hosmer-Lemeshow Chi-square statistic and p-value (generally reported as Pr>ChiSq). If a small p-value (often less than 0.05) indicates a poor model fit with data, the difference between observed and expected frequencies is exceedingly high. However, a big p-value does not indicate a good model fit; all it means is a lack of adequate evidence to reject the null hypothesis. This ambiguity, especially under low levels of statistical power, is one of the largest criticisms of the H-L test (Fagerland and Hosmer, 2012).

Chapter Four - Data Description and Analysis

4.1 Descriptive Statistics and Data Characteristics

This section gives an in-depth analysis on the composition of the dataset and the statistical properties of major finance ratios through a span of 5 years. The analysis is concerned with establishing patterns in the way sample distributions and ratios behave that can act as such indicators of financial distress which are useful for explanatory modeling in the future.

4.1.1 Temporal Distribution and Sample Composition

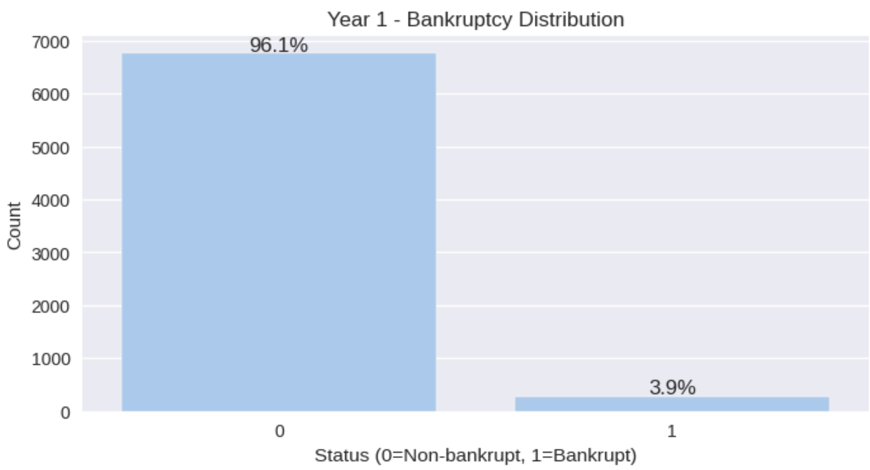

Figure 4.1 - Year 1 bankruptcy distribution

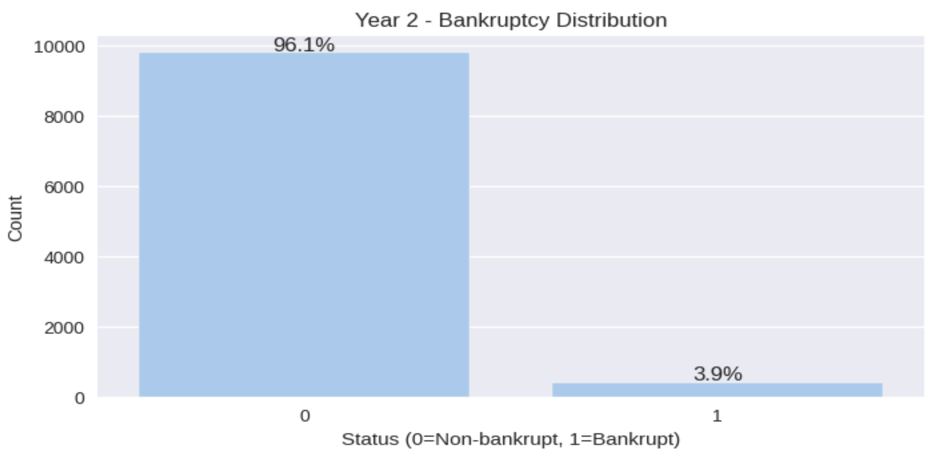

Figure 4.2 - Year 2 bankruptcy distribution

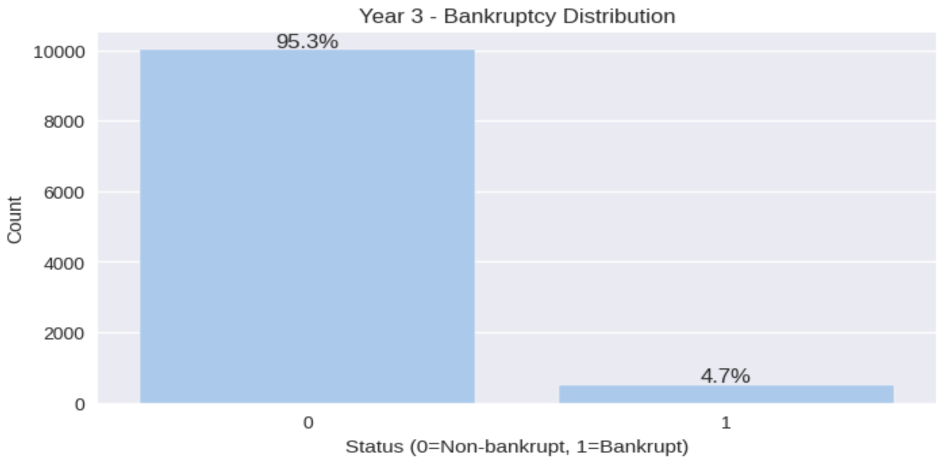

Figure 4.3 - Year 3 bankruptcy distribution

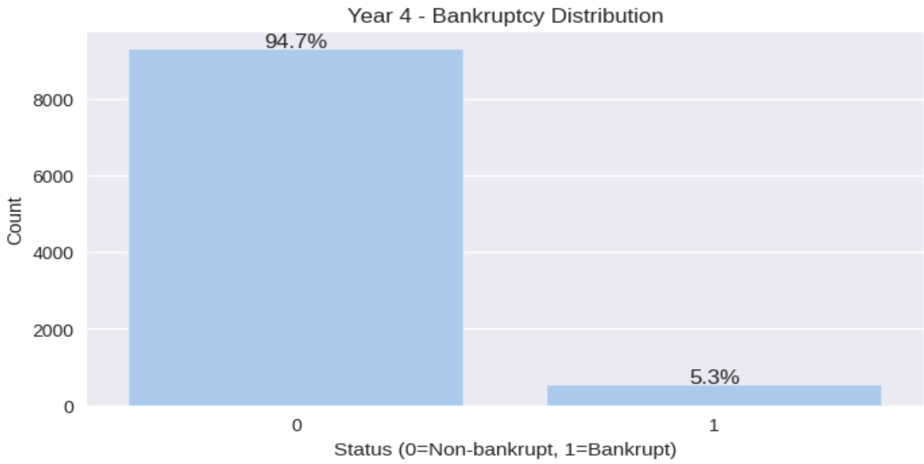

Figure 4.4 - Year 4 bankruptcy distribution

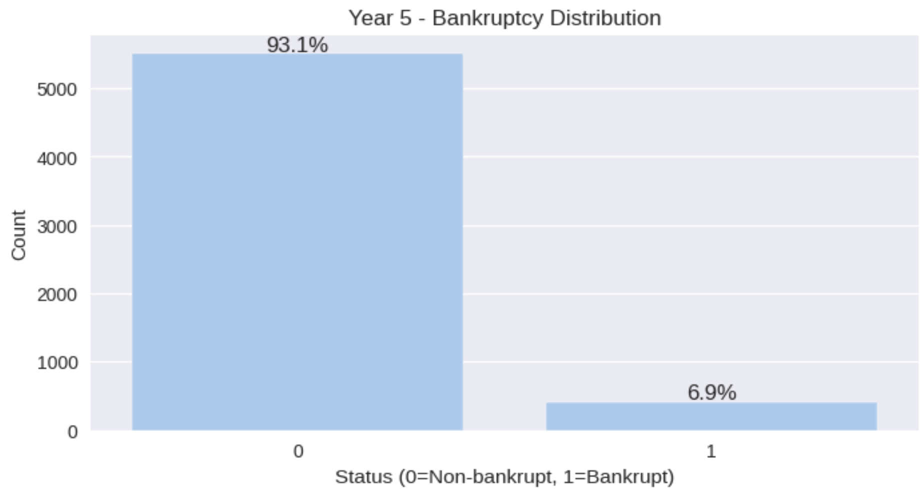

Figure 4.5 - Year 5 bankruptcy distribution

According to the dataset, there is a large variation in the size of the sample with changes along the observation window. The first year has an observation of 7027 firm-year observations which is the smallest sample size. This number increases significantly in Years 2 and 3, reaching 10,173 and 10,503 observations in respective years, and then declining to 9,792 in Year 4 and 5,910 in the last year. This pattern of attrition might be illustration of normal business cycles, where firms are leaving the dataset as a result of bankruptcy, mergers, or other market forces. The decreasing sample over time in the latter years also bring up key considerations with any possible survivorship bias that need to be accounted for in future analyses

Each year's financial data was cleaned, indefinite values were addressed, low-variance and highly correlated features were removed, multicollinearity was handled using VIF, and numeric columns were scaled using robust methods. The cleaned data allowed a clearer view of the most distinctive features separating bankrupt from non-bankrupt firms over a five-year period.

4.1.2 Financial Ratio Distribution Analysis

Some of the financial ratios like Attr5 (defensive interval period) and Attr6 (retained earnings to total assets) have negative means in Year 4 and this can be interpreted to indicate that a large number of firms in the sample had poor financial results in those years. As an illustration, the analysis output indicates that Year 4 experienced the negative mean of Attr6 (retained earnings to total assets) which is the manifestation of the prevalence of unfavorable conditions. All the calculations with yearly means, median, standard deviations, and standard deviation to mean ratios were directly obtained out of the data using Python scripts.

Table 4.1- Year 1 statistical

	count	mean	std	min
Attr5	7027	-0.15	5.56	-36.92
Attr6	7027	0.48	1.45	-4.93
Attr13	7027	0.24	1.14	-2.46
Attr15	7027	0.4	2.76	-11.72
Attr27	7027	10.41	66.1	-2.06
Attr28	7027	1.04	4.03	-1.58
Attr37	7027	5.13	22.02	-1.24
Attr41	7027	0.79	4.15	-9.97
Attr45	7027	1.33	5.69	-2.96
Attr55	7027	0.94	4.03	-7.45
Attr56	7027	0.3	1.26	-2.62
Attr57	7027	0.18	1.23	-4.5
Attr59	7027	1.02	2.93	-3.66
Attr60	7027	1.4	5.61	-0.61
Attr61	7027	0.53	1.96	-0.96

Table 4.2 - Year two statistical

	count	mean	std	min
Attr1	10173	0.11	1.27	-4.63
Attr4	10173	0.78	2.61	-0.76
Attr5	10173	0.24	6.35	-32.75
Attr6	10173	0.24	3.91	-21.98
Attr12	10173	0.69	2.57	-3.78
Attr13	10173	0.13	1.44	-5.68
Attr15	10173	0.24	3.53	-16.77
Attr20	10173	0.26	0.97	-0.69
Attr27	10173	26.62	178.83	-30.17
Attr28	10173	1.42	5.91	-2.96
Attr30	10173	0.35	1.66	-2.36
Attr37	10173	9.14	40.97	-2.25
Attr39	10173	0.06	1.49	-7.8
Attr41	10173	0.89	5.87	-13.44
Attr45	10173	1.41	7.43	-9.83
Attr55	10173	1.05	4.92	-10.94

Attr56	10173	0.23	1.58	-5.82
Attr57	10173	0.07	2.33	-12.51
Attr59	10173	1.03	4.1	-15.7
Attr60	10173	1.59	6.49	-0.6
Attr61	10173	0.64	2.31	-1.05
Attr64	10173	1.73	7.1	-0.55

Table 4.3 - Year three statistical

	count	mean	std	min
Attr1	10503	0.06	1.43	-5.86
Attr4	10503	0.79	2.7	-0.74
Attr5	10503	0.03	5.47	-29.74
Attr6	10503	0	4.85	-28.91
Attr12	10503	0.63	2.67	-5.25
Attr13	10503	0.1	1.53	-6.63
Attr15	10503	0.41	3.53	-13.77
Attr20	10503	0.29	1.04	-0.7
Attr24	10503	0.13	1.11	-3.65
Attr27	10503	22.91	161.97	-79.79
Attr28	10503	1.39	5.56	-3.32
Attr30	10503	0.4	1.96	-2.45
Attr37	10503	13.42	62.73	-2.6
Attr39	10503	-0.04	1.67	-9.14
Attr41	10503	0.9	6.04	-14.61
Attr45	10503	1.17	7.01	-15.22
Attr55	10503	1.02	4.44	-8.98
Attr56	10503	0.18	1.68	-6.25
Attr57	10503	0.09	2.36	-12.09
Attr59	10503	1.33	4.82	-11.27
Attr60	10503	1.65	7.23	-0.62
Attr61	10503	0.65	2.33	-1.01
Attr64	10503	1.72	6.65	-0.54

Table 4.4 - Year four statistical

	count	mean	std	min
Attr1	9792	0.04	1.45	-5.83
Attr3	9792	-0.01	0.9	-3.81
Attr4	9792	0.86	3.11	-0.79
Attr5	9792	-0.01	5.66	-32.48
Attr6	9792	-0.33	5.82	-34.26
Attr12	9792	0.67	2.86	-5.42
Attr13	9792	0.09	1.72	-8.37
Attr15	9792	0.31	3.05	-12.17
Attr20	9792	0.29	1.09	-0.71
Attr24	9792	0.08	1.23	-4.55
Attr27	9792	22.37	161.05	-87.69
Attr28	9792	1.49	6.19	-3.95
Attr30	9792	0.46	2.38	-3.01
Attr32	9792	0.52	2.14	-0.9
Attr37	9792	13.26	61.51	-2.6
Attr39	9792	-0.14	2.26	-15
Attr41	9792	0.73	5.32	-14.04
Attr45	9792	1.23	7.94	-15.55
Attr55	9792	1.11	4.56	-8.22
Attr56	9792	0.15	2.01	-9.86
Attr57	9792	0.04	2.61	-15.13
Attr59	9792	1.22	4.35	-12.56
Attr60	9792	1.66	7.36	-0.66
Attr61	9792	0.66	2.48	-1.03
Attr64	9792	1.71	6.5	-0.56

Table 4.5 - Year 5 statistical

	count	mean	std	min
Attr5	5910	0.01	1.68	-7.67
Attr6	5910	0.03	3.29	-19.85
Attr12	5910	0.57	2.13	-3.59
Attr13	5910	0.11	1.35	-5.1
Attr15	5910	0.37	3.09	-11.14
Attr24	5910	0.08	1.16	-4.01
Attr27	5910	20.85	139.02	-23.54
Attr28	5910	1.57	6.88	-3.47

Attr30	5910	0.29	1.59	-3.09
Attr37	5910	10.61	47.85	-1.8
Attr39	5910	-0.01	1.6	-9.27
Attr41	5910	0.9	5.66	-11.79
Attr45	5910	0.88	6.07	-16.72
Attr55	5910	0.93	3.59	-6.33
Attr56	5910	0.22	1.65	-6.02
Attr57	5910	0.02	2.49	-13.9
Attr59	5910	1.22	4.19	-10.28
Attr60	5910	1.53	6.99	-0.69
Attr61	5910	0.54	2.07	-1.04

The Most extreme volatility is seen in Attr27 (interest coverage) in Year 2, where the calculated standard deviation 178.83 while the mean is only 26.61. This was observed directly in the descriptive summary statistics. Such a high level of dispersion indicates the presence of extreme values that could skew model training unless properly treated. When examining central tendency, the analysis shows evidence of skewness in several variables. For instance, in Year 4, Attr27 (interest coverage) has a mean of 22.36 and a median of 0.00, suggesting a strongly right-skewed distribution. This large gap implies that while most firms had low values, a few had extremely high values. To address this, winsorization (*data transformation technique used to limit the effect of extreme outliers by replacing the smallest and largest values in a dataset with the closest values within a specified percentile range.*) was applied during preprocessing to reduce the influence of outliers without completely removing them.

To evaluate the importance and determinant strength of each attribute, statistical tests were conducted. In particular, p-values were computed using probit and logistic regression models to assess whether differences in the attribute distributions significantly corresponded to the target outcome. Attributes such as Attr6 (retained earnings to total assets), Attr13 (adjusted gross margin), Attr41 (debt payback period) had p-values less than 0.05 especially Attr55 (working capital) dominating in all year in both models indicating that their association with the target variable was statistically significant. This confirms that these attributes are not only statistically relevant but may also serve as key variables in model development.

4.2 Variable Selection and Model Preparation

The process of variable selection and model preparation involved rigorous statistical analysis to identify and retain the most relevant financial ratios while eliminating redundant or problematic features. This was vital so that the important features from the data would be found in the models and would also remain accurate and understandable for interpretation.

In the preprocessing, it was found that a significant group of variables were highly correlated. In every year, between 31 and 33 features showed strong, high correlations with each other. Various variables such as those labeled Attr7 (return on asset-ROA) through Attr11 (gross profit return on asset) (various liquidity and profitability ratios), Attr14 (liabilities turnover period), and many in the Attr16 (cashflow to total liabilities) to Attr63 (short term liability turnover) group (comprising various financial structure, turnover, and performance metrics), were included initially. They were later omitted because having them would bring noise and overlapping variables, without contributing new knowledge. The persistence of these signals over several years supports the idea that some financial ratios stay connected.

To address multicollinearity, a Variance Inflation Factor (VIF) analysis was conducted. Variables with a VIF greater than 5 were excluded, as they exhibited high correlation with other features, which could distort the model’s accuracy. While reducing the number of attributes raised justified concerns, it was necessary to enhance the model's performance and improve interpretability. The remaining attributes were selected based on their low multicollinearity and strong individual explanatory power. Statistical measures, including differences in mean, median, and variance, along with machine learning-based feature selection techniques; such as correlation analysis and recursive feature elimination, guided this selection process.

Interest Coverage (Attr27), quick assets to long-term liabilities (Attr37), retained earnings to total assets (Attr6), and net profit to inventory (Attr45) consistently stood out across all five years due to their distinctive statistical behavior. Attr27 (interest coverage) and Attr37 (quick assets to long-term liabilities) showed the large absolute change in mean value from Year 1 to Year 5 indicating high sensitivity to financial distress. Attr6 (retained earnings to total assets) demonstrated a steady decline in mean and skewness over time, reflecting the erosion of retained earnings in bankrupt firms. Attr45 (net profit to inventory) and Attr55 ((working capital), maintained consistently high positive skew across all years, suggesting that while most firms performed poorly, a subset (likely non-bankrupt) had significantly higher profitability relative to inventory. These persistent patterns made them effective indicators for distinguishing between bankrupt and non-bankrupt firms.

Throughout the annual evaluation, a consistent pattern of feature selection emerged across the five years. In Year 1, 15 features were retained in the models, this trend continued into Year 2, where 22 features were kept, by the 3rd year, 23 were maintained, the selection process in Year 4, 25 features were kept Finally, in Year 5, 19 features were retained.

4.3 Logit Regression Analysis: Initial and Filtered Model Comparison

This study examines the performance of logit regression models in explaining corporate bankruptcy across five years, comparing initial models (with all variables) against filtered models (retaining only statistically significant Explainers at $p < 0.05$). The analysis evaluates model fit, explanatory power, and economic interpretation to identify robust bankruptcy indicators.

Table 4.6 - Logit model comparison

Year	Model	Pseudo R-squared	AIC	BIC	AUC	HL p-value	Accuracy	Num Explainers
1	Initial	0.0784	2148.2	2257.923	0.7297	0.11061872	0.96143	15
1	Filtered	0.0745	2136.73	2177.876	0.7313	0.02173888	0.96143	5
2	Initial	0.0402	3282.99	3449.232	0.681	1.0149E-05	0.9606	22
2	Filtered	0.0336	3273.21	3323.812	0.6688	0.00049139	0.96068	6
3	Initial	0.0617	3792.33	3966.556	0.7249	6.3851E-10	0.9529	23
3	Filtered	0.057	3780.83	3846.174	0.7198	1.7852E-09	0.95277	8
4	Initial	0.081	3761.02	3947.941	0.7327	5.7236E-07	0.94669	25
4	Filtered	0.0755	3751.02	3822.916	0.7291	5.8855E-06	0.94679	9
5	Initial	0.2025	2415.45	2549.139	0.8163	0.00107207	0.93248	19
5	Filtered	0.1966	2412.91	2479.758	0.8183	5.2624E-05	0.93096	9

Table 4.7 - All significant explainer

Year	Explainer	Coefficient	P-value - initial	P-value - Filtered	VIF - Initial	VIF - Filtered
1	Attr6	-0.15000954	0.001794648	0.001902237	1.17073444	1.14709449
1	Attr13	-0.68841594	9.93128E-14	1.12379E-15	2.02503723	1.10801865
1	Attr15	0.058328227	0.00048808	0.000660991	1.07574011	1.05797182
1	Attr41	-0.04285831	0.007752765	0.010058013	1.10157535	1.06725712
1	Attr55	-0.06852196	0.012206123	0.011428169	1.09770693	1.06499392
2	Attr6	-0.03618274	0.001387937	0.001077927	1.26231944	1.14650751
2	Attr15	0.038348129	0.002559345	0.002265915	1.02440754	1.0031666
2	Attr28	-0.04293821	0.007020253	0.003962891	2.67474401	2.42434757

2	Attr39	-0.188766	1.2732E-05	6.96567E-12	2.12918352	1.11768604
2	Attr55	-0.04176311	0.008243514	0.008126651	1.09158277	1.04785187
2	Attr64	0.022809287	0.004836684	0.00426213	2.56418873	2.4183993
3	Attr4	-0.0633188	0.023824507	0.014852078	1.67255875	1.05822159
3	Attr13	-0.17315532	2.92393E-05	2.07376E-05	2.42475156	1.96947138
3	Attr15	0.0247765	0.025497357	0.02257355	1.05691102	1.008346
3	Attr24	-0.20381166	0.001240679	5.52816E-05	2.69400625	1.43896266
3	Attr39	-0.1329415	0.000167515	0.000141736	2.35733736	2.06750121
3	Attr55	-0.03976867	0.009918606	0.012906442	1.08370905	1.037727
3	Attr56	0.070789304	0.023869626	0.028851767	1.92004913	1.86225558
3	Attr64	0.01495311	0.022453595	0.007066867	2.36878257	1.00871968
4	Attr1	-0.15785112	0.001646746	4.49912E-05	3.39221284	1.95156697
4	Attr3	-0.21215936	4.12267E-05	0.000133335	2.60834292	1.56196618
4	Attr13	-0.10307612	0.002285027	0.005399151	2.36115529	1.99526645
4	Attr30	-0.08171235	0.000908331	0.000651277	2.125351	1.95423523
4	Attr32	0.081104009	0.008472365	0.000226878	2.06104576	1.77232705
4	Attr39	-0.06491236	0.004372368	0.002816661	2.66414826	1.7828892
4	Attr41	-0.02638957	0.00412014	0.007292091	1.06876729	1.01231047
4	Attr55	-0.04617375	0.012504961	0.006613476	1.10908574	1.07483065
4	Attr61	0.045152732	0.003010799	0.00188203	1.07718994	1.04335085
5	Attr5	-0.11888695	4.91867E-05	0.000374945	1.73674903	1.17174907
5	Attr6	-0.06935017	1.13956E-05	1.78915E-06	1.53897376	1.42808225
5	Attr13	-0.25181986	2.3613E-06	9.99137E-07	2.31807858	1.75013151
5	Attr24	0.248618709	8.1398E-06	8.64443E-06	2.06317074	1.83150368
5	Attr39	-0.29763417	1.48296E-16	2.07181E-19	1.96824042	1.59882585
5	Attr41	-0.04413727	0.00048542	0.000280384	1.05846957	1.00843244
5	Attr45	-0.05443244	0.001371508	0.000650277	1.98291422	1.18943539
5	Attr55	-0.16336638	4.29163E-06	4.653E-06	1.10058894	1.08124243
5	Attr57	-0.04341566	0.047072794	0.006138635	1.16797698	1.09492069

4.4.1 Model Performance and Fit

The first group of models were full, involving all the available financial variables while the second set of models were the filtered ones kept only those having statistically significant coefficients. With the reduction in the number of variables, the models that passed through the filter had highly explanatory power. For example, the pseudo R-squared reduced slightly in Year 1 from 0.078 to 0.075, while in Year 5 it was high at 0.203, therefore indicating that the filtered models retained most of the explanatory power.

The performance of models was also validated by the area under the ROC curve (AUC). The initial and filtered models had almost equal values of AUC with the highest levels of discrimination power (AUC = 0.816 in the initial model, and AUC = 0.818 in the filtered model) being observed in Year 5. This uniformity implies that in the first models, non-significant financial variables did not significantly contribute towards classification accuracy.

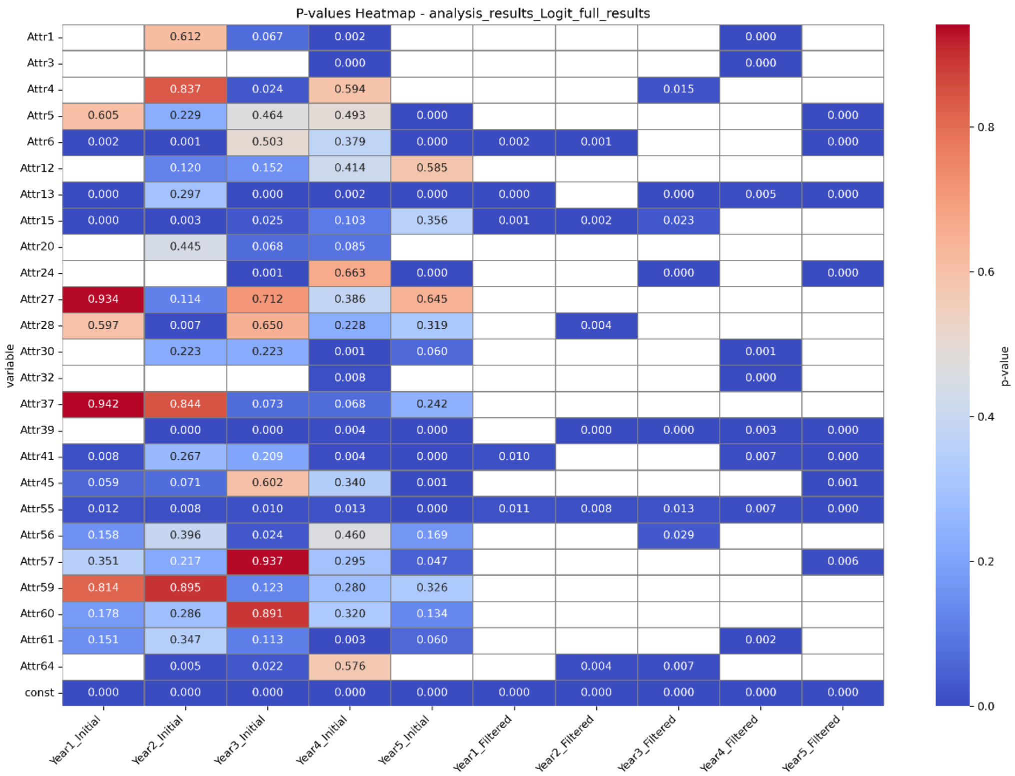

Figure 4.6 - Heatwave Analysis of Logit P-Values

The results of the statistical heatwave logit model show the variation in the importance of financial ratios over the five years. Attr13 (adjusted gross margin), along with Attr55 (working capital) and Attr39 (operating profit margin), showed persistent "hot zones" and had very low p-values in at least four years. For example, Attr13 (adjusted gross margin) always presented a p-value of less than 0.001, which is one of the key indicators of financial stability. Similarly, Attr55 (working capital) and Attr39 (operating profit margin) also portrayed significant and consistent importance, highlighting the key role of liquidity and profitability in distinguishing between healthy and distressed firms.

Some variables, on the other hand, had intermittent heatwaves and were only significant during certain years. Attr4 (current liquidity), for example, was not significant in Years 1, 2, 4, and 5 but was significant in Year 3 (coefficient = -0.0871, p = 0.0149), signifying that current liquidity pressures were especially active in that year. Likewise, the significance of Attr6 (retained earnings to total assets) and Attr15 (debt service capacity ratio) was stronger in the initial years and diminished in subsequent years, implying that leverage and debt-servicing capacity were important early warning signals that faded over time.

Other characteristics, such as Attr24 (gross profit return on assets) and Attr28 (working capital to fixed assets), were meaningful for no more than two years, indicating short-lived spikes in explanatory significance. Meanwhile, some ratios such as Attr12 (gross profit to short-term liabilities), Attr20 (inventory turnover in days), Attr27 (interest coverage), Attr37 (quick assets to long-term liabilities), Attr59 (long-term debt to equity), and Attr60 (inventory turnover) were statistically insignificant at all filtered years, meaning they do not contribute much to explaining bankruptcy in the Logit model.

Finally, the Logit p-value analysis indicates a time trend in which leverage indicators prevailed in the initial years, while profitability and liquidity variables became more prominent over time, depicting a changing financial trend throughout the study period.

4.4.2 Key Explainers of Bankruptcy

Several variables consistently appeared as significant Explainers across multiple

Leverage and Financial Stability (Attr6 - retained earnings to total Assets): This ratio had a negative coefficient in all the years except the Year 4, which suggested that the higher the retained earnings of the companies, and hence the lesser the dependence on external debt, the less likely it was to become bankrupt. This effect was highest during Year 1 (coefficient -0.1516) and although it became a little smaller in each of the successive years; it was still statistically significant. This trend points to the stabilizing role of internally funded capital structures as regards to the survival of firms.

Liquidity and Solvency (Attr15 - debt service capacity ratio (in days), and Attr55 - working capital): All these ratios are used to determine the capacity of a firm to fulfill its commitments using its operating performance and cash flow management. The coefficient of Attr15 (debt service capacity ratio in days) was positive in the initial years, which means that those companies that took more time to pay debts of their working cash flows were at a higher risk of bankruptcy. Attr55 (working capital) provided negative coefficients over all five years (p-values of 0.00-0.02), which shows that the greater working capital is (a strong liquidity and effective management of short-term resources) the less likely it is that the company will experience financial distress.

Operating Profit Margin: Profitability and Operational Efficiency (Attr39 - operating profit margin and Attr13 - adjusted gross margin): Attr13 (adjusted gross margin) recorded negative coefficients during the later years, which proved that firms that have high operating profitability and internal cash generation are less likely to be bankrupt. Attr39 (operating profit margin) value was found to be statistically significant in four out of five years and had negative coefficients (e.g., -0.30 in Year 5) which trend supports the argument that increased profitability and greater operational margins are protective factors against bankruptcy, which is consistent with the expectation that excellent operations increase financial strength and sustainability prospects.

4.4.3 Economic Interpretation of Coefficients

The negative coefficients for variables like Attr6 (retained earnings to total assets) and Attr39(operating profit margin) align with financial theory, where higher liquidity and profitability reduce bankruptcy likelihood. Conversely, some positive coefficients (e.g., Attr15-debt service capacity ratio (in days) in Year 1) may indicate that certain operational efficiencies, while beneficial in stable conditions, could signal risk under financial distress.

4.4.4 Classification Performance and Limitations

While the overall accuracy across all models exceeded 90%, this performance was partly influenced by class imbalance, as bankrupt firms constituted only a small portion of the dataset. This imbalance contributed to a notably low True Positive Rate (TPR), especially in the early years, recording 0% in Year 1 and only improving to 14.6% by Year 5. These percentages were calculated by dividing the number of correctly predicted bankrupt firms (true positives) by the total actual bankrupt firms (true positives plus false negatives) for each year. For example, in Year 1, no bankrupt firms were correctly predicted out of 271 actual bankrupt cases, resulting in a 0% TPR. By Year 5, 60 bankrupt firms were correctly identified out of 410 actual cases, leading to a TPR of approximately 14.6%. This pattern indicates that explaining bankruptcy is particularly difficult during the early stages of financial distress, where warning signals are less pronounced.

Table 4.8 - Confusion Matrices for Model explanations Across Years 1 to 5

Year	Actual	Predicted 0	Predicted 1	True Positive Rate (Recall)
1	**0**	6756	0	1
	1	271	0	0
2	**0**	9773	0	1
	1	400	0	0
3	**0**	10007	1	0.999
	1	493	2	0
4	**0**	9266	11	0.998
	1	511	4	0.007
5	**0**	5451	49	0.991
	1	350	60	0.146

The filtered logit regression models offered a more concise yet comparably effective approach to bankruptcy explanation relative to the initial models. Among the explainers, key financial ratios, specifically those related to liquidity, profitability, and solvency, consistently emerged as significant indicators.

However, the persistently low TPR in the early years emphasizes the limitations of traditional modeling techniques in detecting early signs of bankruptcy. This highlights the potential value of incorporating additional methods, such as ensemble learning techniques or alternative sampling strategies, to improve the sensitivity of early-stage explanations.

Future research should consider exploring dynamic models that can adapt explainer weightings as firms move closer to bankruptcy. Such models could enhance explanatory accuracy during the most critical periods and better support timely decision-making.

Note: *The figure supporting these findings can be observed in the data analysis section, where year-by-year TPR (recall) and model accuracy metrics are shown.*

4.5 Probit Regression Analysis: Key Findings and Comparison with Logit Models

Table 4.9 - Probit model comparison table

Year	Model	Pseudo R-squared	AIC	BIC	AUC	χ^2 P-value	Accuracy	Num Explainers
1	Initial	0.0765	2152.11	2261.835	0.7297	0	0.96143	15
1	Filtered	0.0735	2140.86	2188.869	0.7311	0	0.96143	6
2	Initial	0.0413	3279.42	3445.658	0.6852	0.902143	0.96068	22
2	Filtered	0.0349	3269.09	3319.684	0.6707	0.910677	0.96068	6
3	Initial	0.0643	3781.79	3956.019	0.7294	0.642406	0.95277	23
3	Filtered	0.0594	3769.3	3827.384	0.72456	0.998075	0.95287	7
4	Initial	0.0833	3751.42	3938.349	0.738	0.00601	0.94679	25
4	Filtered	0.0778	3741.68	3813.582	0.7314	0.050905	0.94653	9
5	Initial	0.2059	2405.44	2539.131	0.8215	0	0.9323	19
5	Filtered	0.2034	2396.66	2476.876	0.8225	0	0.9319	11

This study examines probit regression results for bankruptcy explanation across five years, comparing initial models (with all explainers) against filtered models (retaining only significant explainers at $p < 0.05$) similar to the logit regression. The analysis evaluates model performance, key Explainers, and marginal effects while contrasting findings with earlier logit regression outcomes.

Table 4.10 - All significant explainer

Year	Explainer	Coefficient	P-value - initial	P-value - Filtered	VIF - Initial	VIF - Filtered
1	Attr6	-0.08113651	0.000484716	0.000548807	1.170734437	1.14714245
1	Attr13	-0.31063489	7.85899E-12	1.41143E-14	2.025037228	1.18842171
1	Attr15	0.025886336	0.001730841	0.002051862	1.075740107	1.0585755
1	Attr41	-0.02065644	0.006142959	0.006804107	1.101575354	1.06752854
1	Attr45	0.009071693	0.009013552	0.084704377	2.878013056	1.0833191
1	Attr55	-0.03009092	0.013107203	0.011330218	1.097706927	1.0658011
2	Attr6	-0.01897244	0.000702527	0.000553156	1.26231944	1.14650751
2	Attr15	0.016211662	0.004784896	0.005257067	1.024407544	1.0031666
2	Attr28	-0.02121359	0.003207654	0.001842201	2.67474401	2.42434757
2	Attr39	-0.09361485	3.76074E-05	2.596E-11	2.129183518	1.11768604
2	Attr55	-0.01661368	0.012038658	0.011331064	1.091582769	1.04785187
2	Attr64	0.012044614	0.003800817	0.003723454	2.564188732	2.4183993

3	Attr4	-0.02578218	0.033877632	0.016129394	1.67255875	1.05820859
3	Attr13	-0.08421645	8.2819E-06	2.73524E-05	2.424751565	1.73046513
3	Attr15	0.011534964	0.028513533	0.024850183	1.056911018	1.00832594
3	Attr24	-0.10594235	0.000376066	1.29204E-05	2.694006253	1.43432385
3	Attr39	-0.05302705	0.000156627	0.000809342	2.357337358	1.7271577
3	Attr55	-0.01584669	0.016783897	0.019376813	1.083709046	1.0376098
3	Attr64	0.007695358	0.035043245	0.005578504	2.368782572	1.00749577
4	Attr1	-0.07068357	0.003740568	0.00027258	3.392212839	1.95156697
4	Attr3	-0.11346332	4.18309E-05	3.48242E-05	2.608342921	1.56196618
4	Attr13	-0.0573081	0.000274192	0.001200541	2.361155293	1.99526645
4	Attr30	-0.03880849	0.000703811	0.000952212	2.125351001	1.95423523
4	Attr32	0.042742456	0.002849344	0.000135069	2.061045757	1.77232705
4	Attr39	-0.03769288	0.003843346	0.000421684	2.664148256	1.7828892
4	Attr41	-0.01216577	0.003143515	0.005890393	1.068767289	1.01231047
4	Attr55	-0.01779101	0.021730958	0.012604336	1.109085736	1.07483065
4	Attr61	0.024538434	0.002009315	0.000946578	1.077189944	1.04335085
5	Attr5	-0.06643365	5.12188E-05	6.36794E-05	1.736749031	1.18076659
5	Attr6	-0.03624278	7.15453E-06	4.45487E-06	1.538973762	1.44440146
5	Attr13	-0.13764821	4.51447E-07	1.67813E-07	2.318078577	1.78408122
5	Attr24	0.121099936	4.40162E-05	1.95006E-05	2.063170735	1.83440603
5	Attr39	-0.16942341	8.14876E-18	1.14633E-22	1.968240424	1.61075339
5	Attr41	-0.02019923	0.000455855	0.000229107	1.058469572	1.00961391
5	Attr45	-0.02187887	0.000980187	0.000776787	1.982914216	1.93795337
5	Attr55	-0.06065471	0.000130411	0.000116901	1.100588941	1.08215274
5	Attr57	-0.02137069	0.040848136	0.01055415	1.167976985	1.09500031
5	Attr60	0.009393853	0.029749116	0.022273021	1.657465489	1.64593738
5	Attr61	0.024831467	0.044926398	0.026803281	1.081258047	1.02741155

4.5.1 Model Performance and Fit

The probit models demonstrated strong explanatory power, with pseudo R-squared values ranging from 0.04 (Year 2) to 0.206 (Year 5) for initial models. Filtered models retained most explanatory power, with minimal reductions in fit (e.g., Year 1 pseudo R-squared: 0.073 →0.2 in year 5). AUC values remained robust, peaking at 0.82 in Year 5, confirming the models' ability to distinguish between bankrupt and non-bankrupt firms. Classification accuracy exceeded 93% across all years, though true positive rates (TPR) were low in early years (0% in Year 1), reflecting class imbalance challenges.

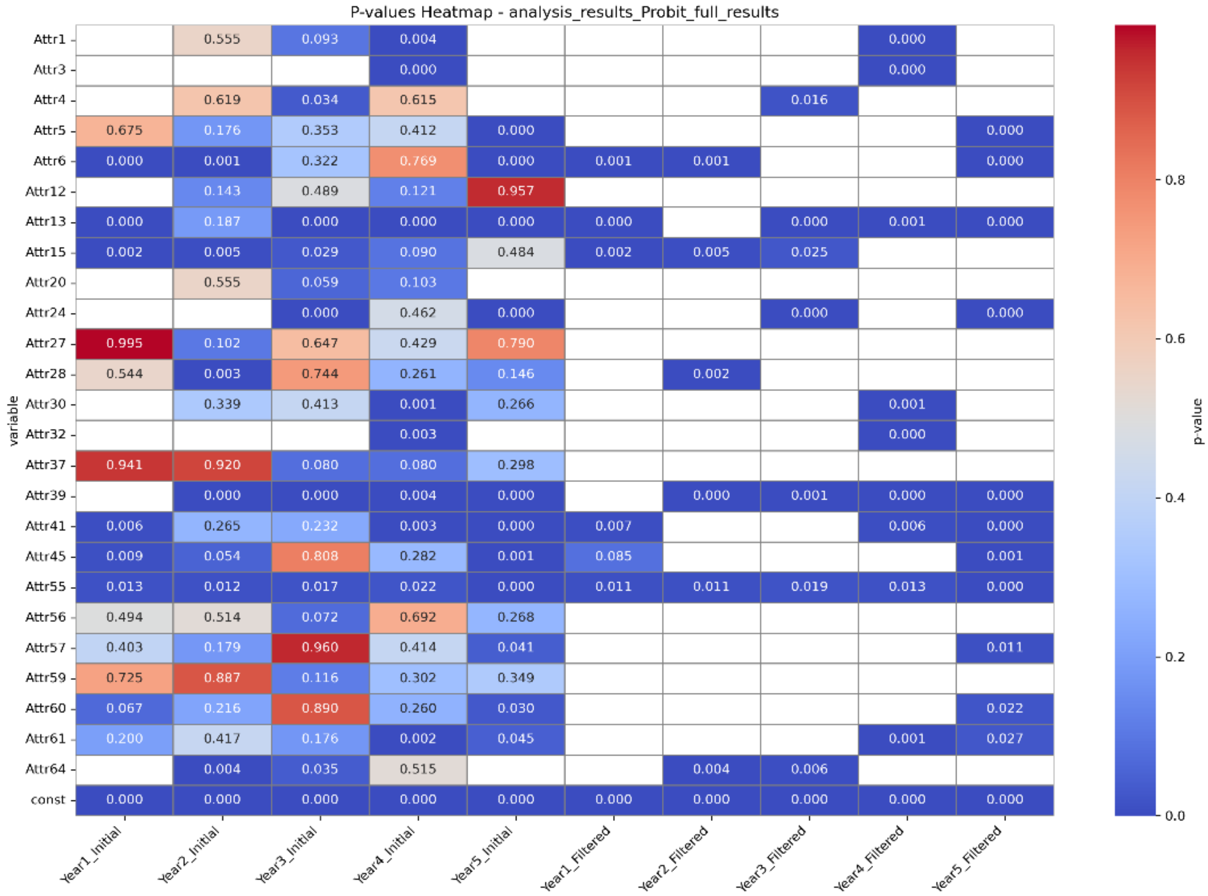

Figure 4.7 - Heatwave Analysis of Probit P-Values

The p-values of the Probit model demonstrate a pattern similar to the statistical "heatwave" observed in the Logit results, although the changes in statistical significance across the five-year period appear smoother. Attr13 (adjusted gross margin) maintains a steady p-value of less than 0.001 in four out of five years, indicating a healthy and stable contribution as a primary discriminant between solvent firms and those that become bankrupt. Attr55 (working capital) remains statistically significant throughout all five years of the filtered period, underscoring the crucial role of working capital as a measure of liquidity and solvency. Profitability indicators, particularly Attr39 (operating profit margin), also retain long-term relevance with consistently low p-values (below 0.001), highlighting the shielding impact of operating profitability on a firm's financial well-being.

There are also attributes whose importance is more intermittent or short-lived, often referred to as temporary heatwaves. For instance, Year 3 (Attr4 - current liquidity) reaches significance (p = 0.0161) but is insignificant in all other years, suggesting a brief period of liquidity stress that temporarily increased bankruptcy risk. Similarly, Attr6 (retained earnings to total assets) and Attr15 (debt service capacity ratio) exhibit higher explanatory in the initial years but lose their ability to predict later on, indicating that leverage and debt coverage served as early warning signs that diminished over time. Other variables, such as Attr1 (return on assets), Attr3 (working capital to total assets), Attr28 (working capital to fixed assets), Attr30 (net liabilities to sales), and Attr32 (accounts payable turnover in days) are situational, as their sensitivity to risk varies depending on contextual factors; they are only significant in one of the five years filtered analyzed.

Conversely, attributes such as Attr12 (gross profit to short-term liabilities), Attr20 (inventory turnover in days), Attr27 (interest coverage), Attr37 (quick assets to long-term liabilities), Attr56 (gross profit margin), and Attr59 (long-term debt to equity) show no statistical significance at any of the filtered years, acting instead as constant "cool zones" with limited explanatory capacity.

Overall, the results of the Probit model reinforce those of the Logit analysis. In the initial years, leverage-related variables, including Attr6 (retained earnings to total assets) and Attr15 (debt service capacity ratio), predominate, whereas in later years, liquidity and profitability variables, such as Attr13 (adjusted gross profit), Attr39 (operating profit margin), and Attr55 (working capital), become dominant. This progression reflects a broader shift in firm financing structures, whereby external debt ceases to serve as the principal explainer of bankruptcy risk, giving way to an increased emphasis on internal performance as the key determinant of financial resilience.

4.5.2 Key Explainers and Economic Interpretation

Consistent with financial theory, liquidity, profitability, solvency and leverage - Attr6 (retained earnings to total Assets), Attr13 (adjusted gross margin), Attr15 (debt service capacity ratio (in days)), Attr39 (operating profit margin), Attr41 (debt payback period) and Attr55 (working capital) emerged as critical variables based on their featuring in minimum of three years similar to logit regression financial variables. Attr6 (retained earnings to total assets) showed a negative coefficient (e.g., -0.15 in Year 1), indicating that higher solvency reduces bankruptcy risk. Attr39 (operating profit margin) exhibited a strong negative effect (Year 5: -0.353), aligning with expectations that profitable firms are less likely to fail.

4.5.4 Comparison with Logit Models

Probit and logit models produced similar results in terms of significant financial variables and directional effects. For example, Attr6 (retained earnings to total assets), Attr39 (operating profit margin), and Attr55 (working capital) were consistently negative in both models. However, probit coefficients were generally smaller in magnitude due to the probit's assumption of a normal distribution (vs. logistic in logit). (Busenbark, 2022) Marginal effects, calculated at mean values, confirmed comparable economic significance, e.g., a one-unit increase in Attr6 (retained earnings to total assets) reduced bankruptcy probability by ~5% in both models for Year 1.

4.5.5 Multicollinearity and Robustness

A Variance Inflation Factor (VIF) test was used to assess the strength of both the Logit and Probit regression models. Regardless of the model year, most VIFs were below 4 (four), indicating that multicollinearity was not a major concern overall. However, some variables recorded moderately high VIF scores. The highest values were observed in Attr1 – return on asset (3.39 in year 4), Attr28 – working capital to fixed asset (2.67 in year 2), Attr45 – net profit to inventory (2.88 in year 1), Attr39 – operating profit margin (2.66 in year 4), and Attr3 – working capital to total assets (2.61 in year 4) in the Probit models. All other financial variables, such as Attr13 – adjusted gross profit (2.02–2.43) and Attr64 – fixed asset turnover (2.24–2.56), remained well within acceptable ranges based on the significant initial and filtered explainers.

The VIF pattern was similar in the Logit models, where only a few attributes exceeded the threshold of four (4), with the highest ones being Attr1 – return on asset (3.39 in year 4), Attr39 – operating profit margin (2.66 in year 4), and Attr3 – working capital to total assets (2.61 in year 4). Most of the explanatory variables, especially Attr6 (retained earnings to total assets), Attr13 (adjusted gross margin), Attr15 (debt service capacity ratio in days), and Attr55 (working capital), consistently recorded VIFs between 1.0 and 2.0, indicating a high level of independence among explainers.

Overall, the assessment of VIF diagnostics in both regression methods verifies that the models are strong and not significantly influenced by multicollinearity. The small elevations observed in certain years do not undermine the credibility of the models but rather indicate expected correlations across related financial indicators. Consequently, both the Logit and Probit estimates can be considered stable and reliable for interpretation.

4.6 Limitations and Recommendations

Although the results were high in accuracy, the low initial true-positive rate suggests how challenging it is to identify bankruptcy at a very early stage. The future studies should include probit or logit specifications with machine-learning algorithms to improve the timeliness and reliability of early warnings. Sensitivity can also be enhanced by including macroeconomic variables to the analytical framework. Their use in the development of focused monitoring systems is supported by the empirical consistency of salient variables when developed using a variety of methodological approaches.

4.7 Model Performance Evaluation

When it comes to testing the models of bankruptcy explanation, both the aspects of statistical validity and practical utility should undergo critical evaluation. This discussion explores the model performance in terms of goodness of fit and classification accuracy and subsequently explores the temporal permanence of explanatory relationships.

4.7.1 Goodness-of-Fit Metrics

The probit regression models showed intra-year stability in the explanatory power. MCFADDEN's pseudo r^2 values were between 0.04 - 0.2, year 2 and 5 respectively therefore moderate but meaningful explanatory relationships. The values compare similarly with those examined from Logistic Regression models, meaning similar explanatory powers between the two methods. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) also trended downwards in value according to time with the lowest value recorded in Year 5 (AIC = 2405; BIC = 2539). This pattern reflects improved efficiency of the model as more data were available and the specifications of the model changed.

(*AIC and BIC are metrics used to select the best model by balancing goodness of fit and model complexity, with lower values indicating a better model*.)

4.7.2 Classification Accuracy

The probit models had excellent discriminatory ability, with the area under the receiver operating characteristics curve (AUC) values being improbably over 0.60 in every year. Year 5 had the highest discriminative ability (AUC = 0.82), whereas the discriminative ability of Year 2 was the lowest (AUC = 0.68). This power of explanation's growth within time probably entails both the collection of more bankruptcy cases and the perfecting of model specifications.Sensitivity analysis demonstrated that the models were very good at explaining true negatives with specificity values presenting 99.3% – 100%. Such high specificity is important for actual applications of risk management, since it reduces false alarms on the firm solvency. True positive rates were more variable with 0% in Year 1 (extreme class imbalance)14.2% in Year 5.

4.8 Stability and Consistency of Key Explanatory Ratios Over Time

4.8.1 Consistency of Significant Ratios

Finding the significance explainers over years indicated considerable regularities in terms of financial ratios that came out as important bankruptcy indicators. Six most important ratios Attr6 (retained earnings to total assets), Attr13 (adjusted gross margin), Attr15 (debt service capacity ratio (in days)), Attr39 (operating profit margin), Attr41 (debt payback period), and Attr55 (working capital) were found to be significant determinants explanatory in at least three of the five annual models. These central variables demonstrated roughly 60-65% of the overall explanatory power in the models of each year. The sustainability of these explanatory relationships in various economic environments increases the faith in the theoretical understanding of such relationships as constituting measuring sticks of financial health.

Importantly, as for all returning explaining variables, the direction of effects was consistent. For example, Attr6 (retained earnings to total assets), Attr55 (working capital) was always negative, and Attr15 (debt service capacity ratio (in days) was always positive. This consistency of the signs of the coefficients increases the possibilities of interpretations of the coefficient as either protective or risk factors (negative or positive coefficients respectively) while explaining bankruptcy.

Chapter Five - Discussion of Key Findings

5.1 Significant Financial Ratios and Comparison with Existing Literature

The empirical investigation showed that a set of financial ratios, such as Retained Earnings to Total Assets (Attr6), Adjusted Gross Margin (Attr13), Debt Service Period (Attr15), Operating Profit Margin (Attr39), Debt Payback Period (Attr41), and Working Capital (Attr55) were always significant to explain bankruptcy results across a number of years. Among these, only Operating Profit Margin (Attr39) was found to be strong, with p-values always close to zero in both the Logit and Probit model between Year 2 and Year 5 indicating a stability of its power to explain throughout the years. Conversely, the temporal variability was observed in certain ratios, such as Current Liquidity (Attr4), which only made significance in Year 3, with p-values of 0.023 (Logit) and 0.033 (Probit), meaning that liquidity indicators do not necessarily offer constant explicative values.

These findings align well with the theoretical and empirical outcomes discussed in the literature review. Initial studies like Beaver (1966) and Altman (1968) highlighted the significance of profitability, liquidity and leverage ratios as one of the key indicators of financial decline. These relationships have recently been confirmed through modern research that indicates that such ratios are highly explanatory, regardless of industry or economic setting. The overall importance of profitability indicators (e.g., Adjusted Gross Margin and Operating Profit Margin) is consistent with the arguments by Bhunia and Sarkar (2011) and Mondal and Roy (2013), who have identified a drop-in profitability as a key trigger of financial distress.

In a similar manner, the explanatory value of the liquidity-related measures, especially Working Capital (Attr55), also confirms the hypothesis that a lack of short-term liquidity exposes businesses to the risk of operational interruptions, a fact that is once again reflected in the findings of Ogachi et al. (2020) and Alifiah (2014). The negative coefficients of Retained Earnings to Total Assets (Attr6) again support the significance of accumulated earnings as a buffer against financial shocks as a theoretical assumption formed in the Z-score model by Altman, and later in the empirical evidence. Nonetheless, the poor explanatory power of some ratios, including Gross Profit Margin (Attr56), which has its importance only in Year 3 of the logit model, indicate that explanatory relevance of financial data can be context-dependent, which makes it important to interpret financial indicators carefully.

5.2 Shifting Explanatory Value of Financial Ratio Over Time

Analysing data over time has revealed how different financial ratios contribute to forecasting earnings which can be steady or cannot be trusted in all years. Although a core group of ratios kept explaining or determining outcomes even in the future, other indicators had their explanatory value change over the years. The fixed assets turnover (Attr64) did very well at explaining outcomes in earlier years and appeared in many filtered models, especially in Year 2 and Year 3. Its significance in statistics dropped in the next years, reaching no significance in Year 4 (Logit initial p=0.5764, Probit initial p=0.5149) and not making an appearance in the Year 5 models.

The feature of Attr47 (inventory holding period) mentioned in the original statement can be seen in none of the p-value tables for any of the years considered. So, it is only possible to say that the initial promising explanatory power and later weakened statistical significance are not fully supported or contested by the current set of results.

Nearly all of these trends noticed such as in the case of Attr64 (fixed assets turnover), correspond with changes in the wider economy, as seen with the 2008-2009 financial crisis. When there are economic or rule changes, this can have a strong impact on the meaning and usefulness of certain financial reports. Finally, the results hint that important financial linkages can easily change in significance as the economy and regulatory rules change.

5.3 Model Explanatory Performance

The models have a noticeable increasing tendency of discriminatory power as the firm nears insolvency. That is supported by the fact that the Area Under the Curve (AUC) is increasing to 0.81 in Year 5 after increasing from 0.68 in Year 2, and it is supported by the fact that pseudo R^2 is also correspondingly increasing from 0.0736 in Year 1 to 0.2035 in Year 5, indicating better model specification and stronger explanatory power. The trend shows that the financial fundamentals captured by the ratios would be increasingly more powerful explainer of failure as the distress level increases. An important conclusion has been made concerning the performance characteristics of the models.

They are always near perfect in being specific (99.1% to 100%), correctly determining an overwhelming majority of solvent firms. For example, in Year 5, the model correctly classified 5451 out of 5500 healthy firms, yielding a specificity of 99.1%. Conversely, the critical capacity to detect actual bankruptcies was low at the start, but changed significantly as time progressed to 14.6 % in Year 5, with the model correctly identifying 60 out of 410 actual bankruptcies in terms of sensitivity. This suggests that the model becomes more responsive to financial distress as bankruptcy becomes more imminent. This temporal progression reveals that financial ratios become increasingly informative about bankruptcy risk as companies approach financial distress, with different ratios gaining explanatory power at different stages of decline. The results confirm that while financial ratio analysis maintains strong explanatory capability, overall, the relative importance of specific ratios evolves over time, suggesting the value of dynamic modeling approaches that can adapt to changing economic conditions.

5.2 Theoretical Implications (Logit & Probit Models)

The analysis, employing both Logit and Probit models, strongly validates the financial distress theory of Ohlson (1980). The consistency observed in the identification of significant core explainers of bankruptcy across both modeling approaches, as evidenced by the presented p-values, highlights the robustness of key financial indicators regardless of the underlying distributional assumptions.

Looking at Attr4 (current liquidity), the empirical findings reveal that it has a fluctuating role in the next five years. The variable became statistically significant only on the third year, with a negative correlation with financial vulnerability (Initial coefficient = -0.0258, p = 0.0339; Filtered coefficient = -0.0258, p = 0.0161). During the rest of the periods (Years 1, 2, 4, and 5), Attr4 (current liquidity) was not significant, so it can be concluded that current liquidity was a short-run factor and could only have an effect under specific conditions in that particular year, not on its own.

Conversely, the significance of Attr13 (adjusted gross margin), Attr39 (operating profit margin) and Attr55 (working capital) remained constant throughout most of the years under study. Attr13 (adjusted gross profit) did not lose significance across at least in four years, with p -values of 1.41 x 10 -14 in Year 1 and p -values of 1.68 x 10 -7 in Year 5, indicating a consistent negative association with financial vulnerability. Attr39 (operating profit margin) also continued to have strong relevance in each of the period (e.g., p = 2.6010-1 in Year 2 and 1.1510-2 in Year 5) which highlighted the importance of profitability as a vital safeguarding element against financial hardship. Attr55 (working capital) was moderately significant but consistently significant across all the sample period (e.g., p = 0.0113 in Year 1 and p = 0.0001 in Year 5), thus proving that working capital is a consistent contributor to the financial stability of firms.

In general, although the effect of Attr4 (current liquidity) was neither long-term, the profitability and working capital indicators were steadier in the long run, which makes them more effective financial variables of the ability of a firm to withstand financial stress.

This alignment in identified significant variables across Logit and Probit models suggests that, in practice, the choice between these two link functions may be less consequential for variable selection than previously believed. The theoretical implications are significant, indicating that model selection does not drastically affect which variables emerge as statistically important, provided technical aspects like coefficient scaling (where logit coefficients are approximately 1.6 times larger than probit, a purely technical difference) are acknowledged. While the two models differ slightly in how they handle the tails of the probability distribution, such differences had a negligible impact on which core features were deemed significant in our results. The practically identical marginal effects, typically derived from these models, further reinforce this point. Thus, for bankruptcy explanation, the emphasis should shift away from debating model choice and instead focus on ensuring careful data preparation and robust interpretation of the most consistently significant explainers.

5.3 Research Limitations

5.3.1 Data-Related Constraints

Only quantitative financial information is considered in the analysis, so significant qualitative factors such as management strengths or changes in the industry are not included. Because bankrupt firms make up only a small part of the data, ranging from 3.86% to 6.94% of the total observations across the five-year study period (calculated by dividing the number of bankrupt firms by the total number of firms for each respective year), training the model is still challenging. Since the five-year interval may miss some economic changes, the study cannot fully show how explanatory relationships develop during different periods in the economic cycle.

5.3.2 Methodological Considerations

Logit and probit models are used in the research, but they aren't as good as machine learning at handling more complex relationships found in datasets. The process of using financial data that is just one year old may reduce the power of explanatory analysis in real time. Because of the way we chose features, certain useful variables may have been left out because their correlation was too high.

5.3.3 Generalizability Issues

Focus on Polish manufacturing firms limits the results' applicability. Structural and financial characteristics unique to this sector and country may not reflect conditions in other industries or regions, reducing the broader relevance of the conclusions. Since findings are from a particular historical sample, they may only apply to other settings in rare cases. Because firms may account for things differently, the study may not accurately account for the variations found in the ratios. Moreover, these models do not identify between bankruptcy types (for example, liquidation compared to reorganization) and so could cover up the real differences.

5.3.4 Implementation Challenges

Rather than making final explanations, the models offer probability numbers that need careful decision thresholds to be applied. Recommendations for recalibration at the ideal frequency are not made by the research. Additionally, the possibilities for strategic reporting by companies using explanatory ratios are not explored.

Such restrictions point to new opportunities for research and explain the meaning of the findings now. The findings give a sense of direction, but practical decisions must be made cautiously considering these guidelines. Further research may overcome these issues by including alternative ways to obtain data, trying more sophisticated models and assessing the approach in all kinds of economic periods.

5.4 Practical Recommendations

For Auditors:

In order to maximise monitoring, concentrate on the most elucidate financial ratios. The analysis allows finding operating profit margin (Attr39) to be the most anticipated early-warning indicator, and it is always important in both models and periods. Although current liquidity (Attr4), and gross profit margin (Attr56) can be useful measures of financial health, their explanatory power for bankruptcy risk was inconsistent between the two metrics. Thus, use a form of tiered review processes. Since it is very reliable, an annual reduced operating profit margin (Attr39) of more than 15% should raise an alarm and initiate thorough revenue and expense checks.

In the case of current liquidity (Attr4), a value of less than 1.0 in two quarters in a row, should still result in more rigorous testing of the working capital, as it is an important liquidity measure. Gross profit margin (Attr56) has less consistent explanatory power and therefore should only be used as an auxiliary tool with major declines triggering an analysis of cost of goods sold and pricing policy instead of explicit bankruptcy warning. Audit workpapers should diligently track these key ratios across multiple periods, identifying deteriorating trends early and adjusting the intensity of review procedures based on each ratio's proven explanatory consistency and predefined thresholds.

For Regulators:

Internationalize dynamic monitoring structures, which will automatically change ratio thresholds, in accordance with current economic conditions. In times of recession, put a more emphasis on the liquidity ratios such as the current ratio whereas in expansionary times, put more emphasis on the profitability ratios. Adopt a gradual strategy where regulation measures increase gradually with the levels of companies going out of danger. Instead of having ratio requirements that are fixed, develop industry specific percentiles that are consistent between business cycles. As an example, enforce corrective action plans in instances where the firm operating margin is below the 25th percentile of the industry peer group in three consecutive quarters.

These ratio-based monitoring systems should be integrated into the continuous assessment procedures by both the auditors and the regulators than using periodic reviews alone. Dashboards can utilize real-time data feeds provided by corporate accounting systems, and automatically raise an alert on developing patterns of risks, in accordance with the existing established explanatory relationships. Such a proactive methodology will allow taking of action at an earlier stage and more independent mitigation of risks than the traditional backward-looking analysis. The proposed framework has reasonable flexibility to provide industry-related changes and still retain the explanatory strength of the identified financial ratios.

Chapter Six - Conclusion and Recommendation

Introduction

This paper had attempted to perform a research to find the contribution of accounting financial ratios in bankruptcy explanation and with time-dynamic statistical methods. Bankruptcy is one of the issues of major concern to financial institutions, investors and policymakers, as its impact can be fatal to the stakeholders and the whole economy. Good analysis of bankruptcy risk assists banks and financial institutions to predict early financial distress, financial distress rating, and strategy determination of credit risk. To this effect this study has been proposed with a stringent model constructed by the use of measurement of historical financial ratios over some years to determine reliable determinants of bankruptcy.

6.1 Purpose of the Study

The main objective of the study was to determine the most important indicators of accounting and financial ratio, which are considered to be the explainer of bankruptcy. In particular, it tried to investigate the dynamics of these indicators and identify whether there is stable explainer over a period of five (5) financial years. In contrast to the conventional study in which static models were considered, the researcher employed a dynamic panel regression estimation method and performed comparisons on multi-year basis to capture the change in financial health over time.

6.2 Models Used and Key Contributions

Logistic and probit regression was used in this study on a dynamic dataset during the five financial years. The major contributions of the study can be discussed as:

Dynamic Analysis: Earlier works were mainly viewed through a static model. The study is also in a unique position to compare financial ratios over the five successive years, thereby noting changing trends in the bankruptcy indicator.

Strong Feature Filtering: There was a two-step feature selection process. We first dropped well correlated variables based on correlation matrices and thereafter performed Variance Inflation Factor (VIF). For example, Attr1 (ROA) which has the greatest VIF of 3.392 in Year 4 implying that independence is acceptable.

Consistency Evaluation: The characteristics that were consistent variables during the study. Case in point, the Attr12 (short term liability coverage) exhibited doctrinaire distinction always between the bankrupt and the healthy firms.

Descriptive Clarity: When compared to most previous models, where only coefficients are listed, we added the interpretations and theoretical coherency of the key financial variables, including Attr6 (retained earnings to total asset), which had a negative coefficient, thus matching the explanations of financial theory.

The benefits of these methodological additions are a high degree of predication and interpretability and a reproducible method of modeling in the near future.

6.3 Major Findings

The findings identified a set of financial ratios that consistently emerged as significant Explainers across the five years:

6.3.1 Key Financial Variables

Attr6 (Retained Earnings to Total Asset): showed negative coefficients, showing that the higher ratio of retained earnings to the total assets shows negative bankruptcy risk. This is consistent with the financial theory because retained earnings are indicators of accumulated profits and inbuilt financing power. If an organisation retained earning begins to deflect, explains that bankruptcy is looming.

Attr13 (Adjusted gross profit): This ratio has negative coefficients in the majority of years; thus, it reflects that firms with a higher-adjusted margin are better able to withstand distress because they have more operating surplus and poor adjusted gross profit explain bankruptcy tendency of the firm.

Attr15 (Debt Service Period - Gross Basis): if this ratio is high, this could be an indication of inefficiency in operations. Intriguingly, its coefficient was positive in Year 1 and this seems to suggest that when a firm is stable, then operating expenses might be desirable and when there is a situation of financial stress operating expenses are a sign of financial distress.

Attr39 (Operating Profit Margin): Negative values in this group confirm the relevance that a greater level of profitability in relation to bodies allows escaping financial hammer as it appears in four out of the five years cycle indicating as a major determinant and vice versa.

Attr41 (Debt Payback Period (in months)): This ratio was prominent in three years of the five years dynamic testing, show a strong indicator, it shows the ability of cash flow (profit include depreciation) to cover debt obligations (total liability) in long term.

Attr55 (Working Capital): It is the steadiest ratio and significant over the five years with negative coefficients having to be significant in throughout the 5 years period for both dynamic regressions. An increased working capital indicates greater liquidity and capacity to settle in the short term thus reducing the chances of bankruptcy and vice versa.

Attr61 (Debtors turnover): This turnover ratio has a positive significance in the later years (year 4 and year 5 for probit while in year 4 in the logit) and it means that the timely debts or receivables are collected, the higher the liquidity, but in other cases, aggressive credit practices can reduce its effectiveness.

Attr64 (Fixed Assets turnover): This measure shows how the firm's fixed assets are efficiently utilized to generate sales, and it seemed critical in Year 2 and Year 3 for both regressions. It explains the vitality of the firm to be able to maximize its fixed assets to generate sales. A low fixed asset turnover explains the tendency of the firm toward bankruptcy.

6.4 Recommendations

According to the findings of this research the following is to be recommended:

In the case of Financial Institutions: Introduce dynamic models, in terms of variables like Attr6 - retained earnings to total asset, and Attr39 – operating profit margin in credit scoring models. Such indicators are precursors of the financial health.

Investors: Take a close note of those firms whose working capital (Attr55), (Attr13) adjusted gross profit margin, and retained earnings to total asset (Attr6) are going down as those are possible indicators of bankruptcy.

Policymakers: Give recommendations that could influence firms to report major financial ratios in a common format to help the outside stakeholders to measure bankruptcy risk.

Internal Management: Ratios such as Attr15 (debt service capacity ratio (in days)) and Attr41 (debt payback period) of operational efficiency may reflect on cases where the internal structures of costs are becoming unsustainable at least when the economy is experiencing a weaker phase.

6.5 Future Research Directions

This study is strong and there is a need to pursue further research:

Dataset Expansion: As the next step, it is necessary to implement not only bigger datasets but also fresher ones including post COVID financials because health measures of the financials might have changed significantly.

Industry-Specific Modeling: sector specific heterogeneities may affect the place of financial ratios. As an example, capital intensive industries might require other variables of bankruptcy.

Machine Learning Integration: Logistic regression was overly interpretable, and more explanatory power might be achieved with the ensemble model (Random Forest or Gradient Boosting).

Non-Financial Indicators: The model could be enhanced by including the qualitative indicators such as the quality of governance, the level of market sentiment or credit rating scores.

Geographical Diversity: Inclusion of firms of more countries into the dataset may provide a test of the model generalizability.

6.6 Contribution to Knowledge

This study has evidently added to literature by:

- Showing that certain financial indicators are consistent over time (e.g. Attr6 (retained earnings to total assets))
- Presenting usable models using VIF filtered attributes
- The insertion of a comparison over several years that rarely appears in the pre-existing literature
- Providing useful tips applicable to the banks and investors
- Eliminating the divide between the theoretical assumption and the empirical model behaviour

This study, when compared to the past ones which only employed 1 year of stationary data, can give more elaborate, stronger findings in the corporate bankruptcy explanation.

6.7 Conclusion

In conclusion, this paper confirms the fact that particular accounting and financial ratios may be taken as early warning signals of bankrupted companies. By way of an active examination in various years and element selection, it was established that ratios like, retained earnings to total asset (Attr6), adjusted gross profit (Attr13), working capital (Attr55), and operating profit margin (Attr39) held inordinate roles throughout. Such a model does not only work better in explaining, but also it is also closely be fit with the practical financial theory.

The offered recommendations are applicable by different stakeholders, and the proposed research directions are very promising as the topic of study. Since bankruptcy explanation is essential in an unstable economy, this study provides a bearable, interpretable, and viable model of explaining financial dangers

References

Agrawal, K. and Maheshwari, Y. (2019) 'Efficacy of industry factors for corporate default explanation', *IIMB Management Review*, 31, pp. 71-77.

Alifiah, M.N. (2014) 'explanation of financial distress companies in the trading and services sector in Malaysia using macroeconomic variables', *Procedia - Social and Behavioral Sciences*, 129, pp. 90-98. https://doi.org/10.1016/j.sbspro.2014.03.655

Al-Kassar, T. and Soileau, S.J. (2014) 'Financial performance evaluation and bankruptcy explanation (failure)', *Arab Economic and Business Journal*, 9(2), pp. 147-155.

Altman, E.I. (1968) 'Financial ratios, discriminant analysis and the explanation of corporate bankruptcy', *The Journal of Finance*, 23(4), pp. 589–609. https://doi.org/10.2307/2978933

Araghi, M.K. and Makvandi, S. (2013) 'Comparing logit, probit and multiple discriminant analysis models in explaining of companies', *Asian Journal of Finance & Accounting*, 5(1). https://doi.org/10.5296/ajfa.v5i1.2977

Asif, M., Tiwari, S., Saxena, A., Chaturvedi, S. and Bhardwaj, S. (2024) 'A study of Altman Z-score bankruptcy model for assessing bankruptcy risk of National Stock Exchange-listed companies', *Proceedings on Engineering Sciences*, 6(2), pp. 789–806. https://doi.org/10.24874/PES06.02A.006

Babbie, E. (2013) *The practice of social research*. 13th edn. Belmont, CA: Wadsworth Cengage Learning.

Balasubramanian, S.A., Radhakrishna, G.S., Sridevi, P. and Natarajan, T. (2019) 'Modeling corporate financial distress using financial and non-financial variables: The case of Indian listed companies', *International Journal of Law and Management*, 61(3/4), pp. 457-484. https://doi.org/10.1108/IJLMA-01-2018-0008

Balcaen, S. and 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), pp. 63–93.

Baltagi, B.H. (2005) *Econometric analysis of panel data*. 3rd edn. Chichester: John Wiley & Sons.

Basu, A., Ghosh, A., Mandal, A., Martin, N. and Pardo, L. (2016) 'A Wald-type test statistic for testing linear hypothesis in logistic regression models based on minimum density power divergence estimator', [Journal details missing; please verify].

Beaver, W.H. (1966) ‘Financial ratios as Explainers of failure’, *Journal of Accounting Research*, 4(1), pp. 71–111.

Bellovary, J.L., Giacomino, D.E. and Akers, M.D. (2007) ‘A review of bankruptcy explanation studies: 1930 to present’, *Journal of Financial Education*, pp. 1–42.

Bernstein, S., Colonnelli, E., Giroud, X. and Iverson, B. (2017) *Bankruptcy spillovers*. NBER Working Paper No. 23162. National Bureau of Economic Research. Available at: http://www.nber.org/papers/w23162 [Accessed 13 Apr. 2025].

Bhunia, A. and Sarkar, R. (2011) 'A study of financial distress based on MDA', *Journal of Management Research*, 3(2), pp. 1-11.

Bogdan, W., Boniecki, D., Labaye, E., Marciniak, T. and Nowacki, M. (2015) *Poland 2025: Europe’s new growth engine*. Warsaw: McKinsey & Company.

Bose, I. (2006) ‘Deciding the financial health of dot-coms using rough sets’, *Information & Management*, 43(7), pp. 835–846. https://doi.org/10.1016/j.im.2006.08.001

Brealey, R.A., Myers, S.C. and Marcus, A.J. (2012) *Fundamentals of Corporate Finance*, 7th edn. New York: McGraw-Hill Irwin.

Busenbark, J. R. (2022). A marginal effects approach to interpreting main effects and moderation. Organizational Research Methods, 25, 147–169.

Creswell, J.W. and Creswell, J.D. (2018) *Research design: Qualitative, quantitative, and mixed methods approaches*. 5th edn. Thousand Oaks, CA: SAGE Publications.

Czepiel, S.A. (2002) *Maximum likelihood estimation of logistic regression models: Theory and implementation*. [Unpublished manuscript]. Available at: http://czep.net/stat/mlelr.pdf

Dimitras, A.I., Zanakis, S.H. and Zopounidis, C. (1996) ‘A survey of business failure with an emphasis on explanation methods and industrial application’, *European Journal of Operational Research*, 90(3), pp. 487–513.

Drozdowicz-Bieć, M. (2012) Reasons why Poland avoided the 2007-2009 recession(Working Paper No. 45). Warsaw: RIED, Warsaw School of Economics.

Ezejiofor, R.A. and Okerekeoti, C.U. (2021) ‘Altman bankruptcy explanation model and corporate governance: An empirical study of Nigerian banks’, *International Journal of Trend in Scientific Research and Development*, 5(6), pp. 159–168.

Fagerland, M.W. and Hosmer, D.W. (2012) ‘A generalized Hosmer–Lemeshow goodness-of-fit test for multinomial logistic regression models’, *The Stata Journal*, 12(3), pp. 447–453.

Gajdosikova, D. and Valaskova, K. (2023) 'Bankruptcy explanation model development and its implications on financial performance in Slovakia', *Economics and Culture*, 20(1), pp. 31-42.

Habibi, A. and Iqbal, M. (2021) ‘Benefits of financial ratios for financing sharia banking in Indonesia’, *Jurnal Ekonomi dan Keuangan Syariah*, 5(1), pp. 1–12. https://doi.org/10.29313/amwaluna.v5i1.5299

Hassani, M. and Parsadmehr, N. (2012) 'The presentation of financial crisis forecast pattern (evidence from Tehran Stock Exchange)', *International Journal of Finance and Accounting*, 1(6), pp. 142-147.

Heine, M.L. (2000) *explaining financial distress of companies: Revisiting the Z-Score and ZETA models*. New York: Stern School of Business, New York University.

Horne, J.C. and Wachowicz, J.M. (2009) *Fundamentals of Financial Management*, 13th edn. Singapore: Prentice Hall.

Hosmer, D.W., Hosmer, T. and Lemeshow, S. (1980) 'A Goodness-of-Fit Tests for the Multiple Logistic Regression Model', *Communications in Statistics*, 10, pp. 1043-1069. https://doi.org/10.1080/03610928008827941

Hosmer Jr., D.W., Lemeshow, S. and Sturdivant, R.X. (2013) *Applied logistic regression*, 3rd edn. Hoboken, NJ: John Wiley & Sons. DOI: 10.1002/9781118548387

Hu, B., Shao, J. and Palta, M. (2006) ‘Pseudo-R^2 in logistic regression model’, *Statistica Sinica*, 16(3), pp. 847–860. Available at: https://www.jstor.org/stable/24308289

Hua, Z., Wang, Y., Xu, X., Zhang, B. and Liang, L. (2007) ‘explaining corporate financial distress based on integration of support vector machine and logistic regression’, *Expert Systems with Applications*, 33(2), pp. 434–440. https://doi.org/10.1016/j.eswa.2006.05.006

Jabeur, S.B. (2017) 'Bankruptcy explanation using partial least squares logistic regression', *Journal of Retailing and Consumer Services*, 36, pp. 197-202. https://doi.org/10.1016/j.jretconser.2017.02.006

Jabbar, A.A. (2017) ‘Decision usefulness approach to financial reporting: A case for employees’, *International Journal of Advanced Engineering, Management and Science*, 3(5), pp. 481–487.

Jarvis, P.J. (2024) *A comparison of different methods of bankruptcy explanation*. [No publisher provided].

Jose, A., Philip, M., Prasanna, L.T. and Manjula, M. (2020) 'Comparison of probit and logistic regression models in the analysis of dichotomous outcomes', *Current Research in Biostatistics*, 10, pp. 1–19. DOI: 10.3844/amjbsp.2020.1.19

Kadim, A., Sunardi, N. and Husain, T. (2020) 'The modeling firm's value based on financial ratios, intellectual capital and dividend policy', *Accounting*, 6(5), pp. 859–870.

Keener, M.H. (2013) 'explaining the financial failure of retail companies in the United States', *Journal of Business and Economics Research*, 11(8), pp. 373-380.

Kolay, M., Lemmon, M. and Tashjian, E. (2016) 'Spreading the misery? Sources of bankruptcy spillover in the supply chain', *Journal of Financial and Quantitative Analysis*, 51(6), pp. 1955–1990.

Kordlar, A.E. and Nikbakht, N. (2011) 'Comparing bankruptcy explanation models in Iran', *Business Intelligence Journal*, 4(2), pp. 335–342.

Kusumasari, L., Anggoro, R.W. and Miswanto (2018) 'The effectiveness between emoticons and traditional figures on presenting accounting information', *International Journal of Engineering & Technology*, 7(3.30), pp. 348–350.

Kuvshinov, D., Richter, B. and Zimmermann, K. (2022) *The shifts and the shocks: bank risk, leverage, and the macroeconomy*. ECB Working Paper Series No. 2672. European Central Bank.

Mankin, J.A. and Jewell, J.J. (2014) 'A sorry state of affairs: The problems with financial ratio education', *Academy of Educational Leadership Journal*, 18(4), pp. 195–219.

Maricica, M. and Georgeta, V. (2012) 'Business failure risk analysis using financial ratios', *Procedia - Social and Behavioral Sciences*, 62, pp. 728-732. https://doi.org/10.1016/j.sbspro.2012.09.123

Meeampol, S., Srinammuang, P., Rodpetch, V. and Wongsorntham, A. (2016) 'Comprehensive analysis of bankruptcy explanation on Stock Exchange of Thailand SET 100'.

Miswanto (2012) 'Kebijakan dalam penentuan dan pendanaan modal kerja perusahaan', *Economia*, 8(2).

Miswanto, Kusumasari, L. and Anggoro, R.W. (2018) *Analysis of financial performance with conventional financial ratio and emoticon*. Working Paper.

Mondal, A. and Roy, D. (2013) 'Financial indicators of corporate sickness: A study of Indian steel industry', *South Asian Journal of Management*, 20(2), p. 85.

Murray, L., Nguyen, H., Lee, Y.-F., Remmenga, M.D. and Smith, D.W. (2012) 'Variance inflation factors in regression models with dummy variables', *24th Annual Conference on Applied Statistics in Agriculture: Proceedings*. Manhattan, KS: New Prairie Press, Kansas State University.

Ogachi, D., Ndege, R., Gaturu, P. and Zoltan, Z. (2020) 'Corporate bankruptcy explanation model, a special focus on listed companies in Kenya', *Journal of Risk and Financial Management*, 13(3), p. 47.

Ohlson, J.A. (1980) 'Financial ratios and the probabilistic explanation of bankruptcy', *Journal of Accounting Research*, 18(1), pp. 109–131. https://doi.org/10.2307/2490395

Pietrzak, M. (2022). Can financial sector distress be detected early? *Borsa Istanbul Review*, *22*, 1132–1144

Radovanovic, J. and Haas, C. (2023) 'The evaluation of bankruptcy explanation models based on socio-economic costs', *Expert Systems with Applications*, 227, p. 120275. https://doi.org/10.1016/j.eswa.2023.120275

Rashid, C.A. (2020) 'Balanced Score Card and Benchmarking as an Accounting Tool to Evaluate Morrison's Performance', *Journal of Global Economics and Business*, 1(3), pp. 59-72.

Ross, S.A. (2016) *Corporate Finance*, 11th edn. New York: McGraw-Hill.

Salome, I.O., Ironkwe, U.I. and Akani, F.N. (2021) 'Real activities earnings management and the financial performance of quoted manufacturing companies in Nigeria', *European Scholar Journal*, 2(7), pp. 85-94.

Scott, W.R. (2009) *Financial Accounting Theory*, 5th edn. Toronto: Pearson Prentice Hall.

Sheikh, S. and Yahya, M. (2015) *Bankruptcy explanation: static logit and discrete hazard models incorporating macroeconomic dependencies and industry effects*. Master's thesis. Bergen: Norwegian School of Economics.

Surjanovic, N., Lockhart, R. and Loughin, T.M. (2020) *A Generalized Hosmer-Lemeshow Goodness-of-Fit Test for a Family of Generalized Linear Models*. arXiv. Available at: https://doi.org/10.48550/arXiv.2007.11049

Titman, S., Keown, A.J. and Martin, J.D. (2014) *Financial Management: Principles and Applications*, 12th edn. Pearson.

VenkataRamana, N., Azash, S.M. and Ramakrishnaiah, K. (2012) 'Financial performance and explaining the risk of bankruptcy: A case of selected cement companies in India', *International Journal of Public Administration and Management Research*, 1(1), pp. 40-56.

Wang, T.C. and Chen, Y.H. (2006) 'Applying rough sets theory to corporate credit ratings', *IEEE International Conference on Service Operations and Logistics, and Informatics*, pp. 132-136.

Xu, W., Xiao, Z., Dang, X., Yang, D. and Yang, X. (2014) 'Financial ratio selection for business failure explanation using soft set theory', *Knowledge-Based Systems*, 63, pp. 59-67. https://doi.org/10.1016/j.knosys.2014.03.007

Yap, B.C.F., Mohamad, Z. and Chong, K.-R. (2013) 'A longitudinal and cross-industry study on the stability of financial ratios of Malaysian companies', *Accounting and Finance Research*, 2(3), pp. 45-62.

Yap, C.F., Helmi, M.H.M., Munuswamy, S. and Yap, J.R. (2011) 'The Explanatory abilities of financial ratios in explaining company failure in Malaysia using a classic univariate approach', *Australian Journal of Basic and Applied Sciences*, 5(8), pp. 930-938.

Yap, C.F., Munuswamy, S. and Mohamed, Z.B. (2012) 'Evaluating company failure in Malaysia using financial ratios and logistic regression', *Asian Journal of Finance and Accounting*, 4(1), pp. 330-344. https://doi.org/10.5296/ajfa.v4i1.1556

Yazdanfar, D. and Nilsson, M. (2008) 'The bankruptcy determinants of Swedish SMEs', *Institute for Small Business & Entrepreneurship*, 5-7 November 2008, Belfast, N. Ireland.

Zaini, B.J. and Mahmuddin, M. (2019) 'Classifying firms' performance using data mining approaches', *International Journal of Supply Chain Management*, 8(1), p. 690.

Zięba, M., Tomczak, S.K. and Tomczak, J.M. (2016) 'Ensemble boosted trees with synthetic features generation in application to bankruptcy explanation', *Expert Systems with Applications*, 58, pp. 93-101.

Zmijewski, M.E. (1984) 'Methodological issues related to the estimation of financial distress explanation models', *Journal of Accounting Research*, 22, pp. 59-82.

Zohra, K.F., Mohamed, B., Elhamoud, T., Garaibeh, M., Illhem, A. and Naimi, H. (2015) 'Using financial ratios to predict financial distress of Jordanian industrial firms - "Empirical study using logistic regression"', *Academic Journal of Interdisciplinary Studies*, 4(2), pp. 136-142.

Zubanovic, A. and Ahmeti, L. (2020) *The Explanatory power of financial ratios on bankruptcy: A quantitative study of non-listed limited liability SMEs companies in Sweden*. Master's thesis. Jönköping University.