Louvain School of Management
and Nova School of Business and Economics

Validity of traditional bankruptcy prediction models during crisis periods

Author:
Clara Oberhößel

Master's thesis in the view of getting the degrees:
Master of Science in Finance
Master of Science in Management, Professional Focus in Corporate Finance

Supervisor:
James Thewissen and Irem Demirci

Academic Year 2025-2026

Abstract

This thesis evaluates the out-of-sample predictive validity of three traditional accounting-based bankruptcy prediction models Altman's Z-score (1968), Altman's Z"-score (1983), and Ohlson's O-score (1980), on a sample of 104,367 U.S. public firm-year observations drawn from Compustat over 1998-2022, including 377 bankrupt firms across three macroeconomic crisis regimes and ten Fama-French industry groups, to determine their robustness over time. All three models are applied in their original form and re-estimated on modern training data, with classification performance evaluated on a common holdout sample of 148 bankrupt and 2,960 non-bankrupt firm-years. The results indicate that all three models report substantial performance degradation when applying their original published coefficients, driven by structural changes in firm financial profiles since their original estimation periods. Re-estimation recovers meaningful predictive validity for Ohlson's O-score and Altman Z''-score model specifications, achieving combined errors of 30.4% at C* = 0.051 and 35.2% at C* = 0.107, respectively, but fails to do so for the original Altman Z-score. Coefficient analysis reveals a structural shift in the determinants of bankruptcy with leverage losing its predictive weight while the importance of sustained operating losses increases nearly sixfold. Counterintuitively, model performance does not deteriorate considerably during the Global Financial Crisis or Dot-com bust but varies significantly across Fama-French industry groups with the strongest results observed for Business Equipment at a combined error rate of 25.8%. The findings indicate that the Ohlson variable set is the most temporally robust of the three when re-estimating, that cutoff recalibration offers the most accessible performance gain for practitioners, and that industry heterogeneity underscores the need for sector-specific and dynamically updated prediction frameworks.

Keywords: Bankruptcy prediction; Altman Z-score; Ohlson O-score; financial distress; re-estimation; structural change; crisis periods; classification error; industry heterogeneity; Compustat

Declaration of the use of AI tools

The content of my master's thesis is my original output, even if I used generative AI to help prepare it. I master the theories, tools, and methods that I use for the thesis. I verified that the output presented in the thesis is accurate and valid, and that it contains no hallucination, no false reference, etc. I acknowledge that I master the final output and that I am able to explain it and answer questions about it. I made sure not to provide generative AI tools with access to confidential data or information.

By signing this declaration, I affirm that the content of this master's thesis reflects my original work, augmented by the responsible use of AI.

Clara Oberhoessel

Table of Contents

1. Introduction

Losses resulting from a firm's failure can be up to 20% of the bankrupt firm's pre-distress value (Bris et al., 2005). However, the true cost extends well beyond the firm itself, as employees lose their jobs, suppliers absorb unpaid invoices, creditors recognise write-downs, and shareholders' investments are wiped out. While corporate bankruptcy is difficult to predict, it carries disproportionately large consequences, making it a topic of interest to a broad range of stakeholders. The ability to identify firms at risk of failure one year in advance is therefore essential internally to ensure financial stability, and externally for credit allocation or portfolio management, which explains why bankruptcy prediction models are continuously researched.

Despite the development of increasingly sophisticated prediction approaches, such as hazard models or machine-learning techniques, the traditional accounting-based models developed by Altman (1968) and Ohlson (1980) remain the benchmark for academic papers and a first-stage screening tool in practice (Altman, 2018). Both models only require five to nine financial ratios that are easily extractable from any standard financial statement, producing a single interpretable score that can be communicated to non-technical stakeholders and link balance sheet evolutions directly to failure risk. This transparent, ratio-based explanation of why a given firm has been flagged represents an advantage that more recent and complex 'black-box' models cannot replicate. However, Altman's model was calibrated on U.S. manufacturing firms from 1946 to 1965, and Ohlson's on industrial firms from 1970 to 1976, both populations known for capital-intensive operations and primarily tangible asset bases. The changed modern environment raises concern about their usability today. This study reveals that structural changes in financing conditions and the rise of intangible-intensive firms may have shifted the relationships between accounting ratios for bankrupt and non-bankrupt firms. Together, reducing the calibration accuracy and discriminatory validity of traditional models developed for a fundamentally different economic environment. This motivates the re-estimation of the models on modern data and across sectors.

Next to overall structural changes, the 21st century has been affected by three major crisis events, the Dot-com bust, the Great Financial Crisis, and the COVID-19 pandemic. Each represent a distinct period which were marked by a high number of bankruptcies, for instance more than 55,000 U.S. business filings were recorded in 2009 alone (U.S. Courts, 2009). These crises have stress-tested the predictive validity of any model calibrated under stable conditions, especially, as neither Altman's nor Ohlson's model incorporates macroeconomic variables. Yet the practical demand for bankruptcy prediction is arguably highest during crisis periods, as firm failures may be systematically different from normal times and the consequences being more severe, explaining why this study focuses on testing the models one year prior to a crisis.

The analysis is conducted through a systematic re-estimation and out-of-sample evaluation of the Altman Z-score (1968), the Altman Z"-score (1983), and the Ohlson O-score (1980) on a sample of 104,367 U.S. public firm-year observations drawn from Compustat over the period 1998-2022, encompassing 377 bankrupt firms across three distinct crisis periods and ten Fama-French industry groups. Whilst the loss of out-of-sample predictive validity is not a new finding, as Begley et al. (1996) documented significant coefficient instability when applying both models to data from the 1980s, and crisis events have also been researched previously, for instance in the hospitality industry during COVID-19, this study contributes to the literature in several ways. It simultaneously evaluates three classical specifications on a comprehensive modern common dataset including the full structural changes of the early 21st century, as well as covering three crisis periods, and the cross-section of ten industry groups. It provides evidence that the original published coefficients have lost their discriminatory validity when applied to modern U.S. data, but that re-estimating their coefficients recovers meaningful predictive accuracy, and tracks the structural shift in the drivers responsible for classifying corporate bankruptcy from leverage-dominated to loss-persistence-dominated prediction. In addition, and contrary to expectations, the results show that the re-estimated model performance does not vary systematically across crisis periods but does significantly differ across industries, implying model heterogeneity and raising questions about whether sector-specific model calibration would further improve classification accuracy.

The remainder of this thesis is organised as follows. Section 2 reviews the relevant literature and develops the hypotheses. Section 3 describes the data and sample construction. Section 4 explains the empirical methodology. Section 5 reveals the empirical results. Section 6 discusses the findings and their practical implications.

2. Literature review

Predicting corporate distress and bankruptcy has been a central topic of empirical finance for more than fifty years. The academic and practical interest in this question comes from a wide range of stakeholders, including lenders, equity investors, managers, employees, and regulators, who all have strong incentives to identify early warning signals to protect the stake at risk. As a result, a rich set of bankruptcy prediction models has emerged, of which Altman's Z-score (1968) and Ohlson's O-score (1980) represent the two most widely cited and most extensively used. At the same time, more recent literature has raised important concerns about the predictive stability of these models across different time periods and across industries.

2.1. Traditional Prediction Models as Benchmark Tools

In 1968, Altman introduced the first widely adopted multi-ratio bankruptcy prediction model, the Z-Score (Altman, 1968). Building on the univariate analysis of Beaver (1966), Altman used the multiple discriminant analysis (MDA) framework of Fisher (1936) to identify five financial ratios spanning liquidity, profitability, leverage, solvency, and activity, whose linear combination maximised the between-group variance relative to the within-group variance. The resulting discriminant function is given in equation (1):

$$(1)\ Z = 1.2\,X_1 + 1.4\,X_2 + 3.3\,X_3 + 0.6\,X_4 + 1.0\,X_5$$

where X1 = Working Capital/ Total Assets; X2 = Retained Earnings/ Total Assets; X3 = EBIT/ Total Assets; X4 = Market Value of Equity/ Total Liabilities; X5 = Sales/ Total Assets; Z= Overall Discriminant Score. Altman identified a cutoff value of Z = 2.675 as the threshold that minimises the total number of misclassified firms and reports a classification accuracy of 95% one year before bankruptcy. A first extension widened its application by replacing the market-value variable X4 with its book-value equivalent in the Z'-score, which allowed application to privately held firms for which market valuation is not available. The subsequently developed four-variable Z"-Score (Altman, 1983), shown in equation (2), excludes the activity ratio X5 and adds an intercept, to extend application to more diverse industrial sectors. The Z''-score has proved to be a more robust alternative and often outperforms the original Z-Score (Altman, 2018).

$$(2)\ Z'' = 3.25 + 6.56\,X_1 + 3.26\,X_2 + 6.72\,X_3 + 1.05\,X_4$$

Many later developed models used Altman’s model as starting point and inspiration for further refinements, including Ohlson’s O-score (Ohlson, 1980). Ohlson aimed to address the statistical limitations of Altman’s model by using a conditional logit model instead. He identified four broad firm characteristics that should affect the probability of failure: the firm’s size, its financial structure, performance, and current liquidity. While Ohlson also relies on accounting-based information to build his vector of nine independent variables given in equation (3), he applies a logistic transformation in equation (4) that offers an estimated probability of bankruptcy bounded between 0 and 1:

$$(3)\ O = -1.32 - 0.407\,SIZE + 6.03\,TLTA - 1.43\,WCTA + 0.0757\,CLCA - 1.72\,OENEG - 2.37\,NITA - 1.83\,FUTL + 0.285\,INTWO - 0.521\,CHIN$$

$$(4)\ P(Bankruptcy) = 1\,/\,(1 + \exp\,(-\,O\,))$$

Ohlson (1980) reports that a probability cutoff of P = 0.038 minimises the unweighted sum of Type I and Type II errors on his estimation sample, resulting in predictive performance broadly comparable to that reported by Altman.

2.2. Modern Alternatives

Since Altman’s and Ohlson’s original publications, the literature has introduced a significant number of methodological alternatives for predicting distress, such as machine learning or market-based models. Whilst these developments brought new insights into distress prediction and are often seen to achieve better predictive validity, this thesis focuses on Altman’s and Ohlson’s models for the following three reasons. First, to date, the traditional models remain the baseline for prediction and serve as a benchmark against which new contributions are evaluated, such as the Hillegeist et al. (2004) comparison of Altman’s and Ohlson’s models against a market-based approach using a discrete hazard approach. Therefore, assessing their performance and re-estimating their models on modern data are a needed precondition for a fair evaluation of the real or perceived improvements of more recent alternatives. Hillegeist et al. (2004) comparison showed that models based on the Black-Scholes-Merton (BSM) framework outperform Altman’s or Ohlson’s scores. Yet, Altman Z-score remains statistically significant, as it provides incremental information that complements the market-based approach. This explains their continued widespread use, as evidenced in the 29,900 Google Scholar citations of Altman’s original article and 11,050 citations of Ohlson’s work (Google Scholar, 2026).

Second, as both models rely on accounting-based variables, they allow for better interpretation and a clearer economic understanding which is essential when the models are applied as strategic management or supervision tool. As stated by Altman (2018) in his examination of the evolution of the Z-score over the last 50 years, his model is being used, for instance, by managers, who apply it to simulate how strategic decisions (e.g., asset sales) will impact their credit score (Altman, 2018). The contribution of each ratio to the probability of failure can be read off directly from the coefficient, which is also essential for this paper, as it helps understanding why the weights of relevant ratios vary depending on the type and characteristics of the crisis or industry. Thirdly, accounting information can be extracted from publicly available financial statements that do not need the matched market-data infrastructure required by BSM or hazard approaches. This facilitates the application process, lowers the cost of replication, and expands the range of firms to which these models can be applied to.

Nonetheless, prior out-of-sample studies have documented coefficient instability when traditional models are applied to later periods. Begley et al. (1996), report a significant deterioration in predictive performance when applying the original models to 1980s data.
H1: The Altman Z-score, Z''-score, and Ohlson O-score lose significant discriminatory validity when applied to modern U.S. data from 1998-2022 compared to their original reports.

2.3. Testing Prediction Models over different times and across sectors

Classical distress scores are typically calibrated under normal economic conditions, implicitly assuming that the relationships between accounting variables and failure are stable over time. In reality, bankruptcies and distress can be structural, arrive in waves, and tend to cluster with macroeconomic regimes, which is often outside the control of any single firm. Whilst these models have been long standing, several researchers and previous literature have tested regime dependence of bankruptcy models.

Sung et al. (1999) provide direct evidence on the deterioration of model performance under crisis conditions. As they studied the 1997 Korean financial crisis, they find a significant decrease in accuracy of models built under normal economic conditions when applying them to crisis scenarios (Sung et al., 1999). When observing the dynamics of model change from normal to crisis conditions, in the context of the 1997 Korean economic crisis, the model disclosed that accuracy dropped from 66.7 percent to 36.7 percent. Consequently, when testing the models across the most recent events of the 21st century their performance may decline.

H2: The predictive performance of traditional bankruptcy prediction models declines significantly during crisis periods relative to non-crisis periods, as assessed across the dot-com bubble (2001), the Global Financial Crisis (2007-2009), and the COVID-19 pandemic (2020).

Additionally, Sung et al. claim that financial ratios that tend to be the best at predicting failure can differ across environments. Whereas cash flow to total assets and productivity of capital appear more important in normal times, cash flow to liabilities becomes more important in crisis conditions. This is explained by the fact that solvency ratio measures the firm's instant ability to service their debt in a volatile crisis environment where refinancing tends to be difficult. Camska and Klecka (2020) compared the performance of financial distress prediction models across different economic conditions, including a period of expansion in 2017 and recession in 2012 and conclude that changing economic environments do influence the accuracy (Camska & Klecka, 2020). Implying that the general macroeconomic situation should be considered when interpreting results. Duricova et al. (2025) confirm this finding by analysing models across pre-pandemic, pandemic, and post-pandemic periods, indicating that many models experienced a decline in predictive accuracy during the pandemic period, explained by coefficient instability. Thus, as crisis periods have distinct characteristics, they also may have different impacts on the accuracy of various prediction models.

H3: The magnitude of performance deterioration differs systematically across crisis types and model used due to differences in underlying economic mechanisms.

The effect of industries on the validity of bankruptcy prediction models can be seen in several comparative studies. Candera's (2024) study focuses on service companies and establishes that model performance differs when applied to the same sample and that traditional models underperform during the pandemic, as they include only financial variables and exclude the non-financial information that became more critical during COVID-19. This advocates sector-specific adjustments to increase prediction validity. Huo et al. (2024) compares two approaches for restaurant firms, an industry characterised by high volatility. They find that logistic regression tends to be a more accurate predictor than using a multivariate discriminant analysis. Together, indicating that model performance is sensitive to industry context and that the logit-based Ohlson framework may generalise more effectively across heterogeneous sectors.

H4: The Ohlson O-score generalises more effectively across industries than the Altman Z- or Z''-score models.

2.4. Case study on Dot-com Bubble, Great Financial Crisis, and COVID-19 Pandemic

The 21st century has been affected by three major events that impacted the macro economical dynamics. The first one being the dot-com bubble, that lasted from March 2001 until November 2001, which was a tech-heavy bubble. During a rise in popularity of the internet sector in the late 1990s and early 2000s, high investments were made into this market, equalling to 6% of the market capitalisation of all US public companies. At the same time, many firms were characterised by weak profits and cash burn, whilst still retaining a stock price that did not reflect the actual underlying economic activities of many, resulting in overvaluation (Hirakubo & Friedman, 2002). Consequently, more than 210 Dot-com companies failed in 2000 (Corry, 2001). This impacted the relevance of traditional accounting information usually used to assess a firm. Thus, the relationship between market value and their actual traditional accounting information vastly decreased, and profitability and liquidity signals dominated over traditional solvency signals (Morris & Alam, 2008).

The second crisis faced was the Great Financial Crisis (GFC) of 2008-2009, which was a severe global economic downturn caused by a mortgage crisis. Even though all industries were impacted, the industries experiencing the most affect by it were Airlines, Diversified Financials, Motor Vehicles and Parts industries (Dzikowska & Jankowska, 2012). Key characteristics included differences in return on revenues, return on assets, and profit per employee. Männasoon et al. (2018) detect that companies showing a high investment intensity and debt financing were the ones most likely to experience financial distress (Männasoo et al., 2018).

Lastly, the COVID-19 crisis was a global economic shock triggered by a health pandemic from January 2020 until March 2020 (Tasley, 2023). It was characterised by a sudden economic disruption and systemic financial challenges, such as high levels of pre-existing debt and risks of loan non-compliance. Though it was a worldwide crisis, affecting almost every industry, the most impact has been seen on tourism, transport, agriculture, and mining. When examining firms across 24 countries, Bozkurt & Kaya reveal that firms with problems in accessing finance, younger firms, and more indebted firms had increased level of bankruptcy risk (Bozkurt & Kaya, 2023). Nonetheless, the study discovered that the primary factors causing businesses to face financial distress and bankruptcy across during COVID-19 were country-specific features, such as governance quality or shutdown duration.

Each of these crises originate from distinct roots and affects financial ratios differently, suggesting that the predictive accuracy of traditional bankruptcy models, which rely solely on accounting variables without macroeconomic inputs, may vary systematically across crisis types. This thesis uses these three events to test whether predictive accuracy deteriorates during crisis periods and whether the type of crisis matters for the magnitude of any such deterioration.

2.5. Research Question and Contribution

The preceding review establishes that even though traditional bankruptcy prediction models remain relevant as a benchmark and in practical applications, due to their interpretability, simplicity, and comparatively low data requirement, the existing literature raises three well-documented concerns: the deterioration of out-of-sample performance over time (Begley et al., 1996), the potential sensitivity of model accuracy to macroeconomic regime (Sung et al., 1999; Camska and Klecka, 2020; Duricova et al., 2025), and the potential for cross-industry variation in predictive validity (Huo et al., 2024; Candera, 2024). However, no existing study simultaneously examines all three concerns across all three traditional model specifications on a comprehensive modern U.S. dataset spanning 21st century structural changes, three distinct crisis regimes, and the full cross-section of Fama-French’s ten industry groups.

This thesis addresses that gap by testing and re-estimating the Altman Z-score (1968), the Altman Z"-score (1983), and the Ohlson O-score (1980) on U.S. Compustat data from 1998 to 2022. It follows the same statistical methodology as the original authors and evaluates whether the original and the re-estimated model specifications produce a relevant classification performance on the modern holdout sample. This is done to assess whether the underlying variable combinations retain validity even if the coefficients require updating.

H5: Re-estimating traditional bankruptcy prediction models on modern data improves predictive performance relative to the original published model specifications.

Additionally, whilst previous studies evaluate model accuracy for one specific crisis event e.g. COVID, one specific industry e.g. Gastronomy, or a specific geography e.g. Korea, this thesis contributes by simultaneously addressing all three dimensions on a single, common dataset. In doing so, it provides an updated assessment of whether the Altman and Ohlson models remain suitable for usage in the modern era, identifies the structural drivers behind performance changes, and provides practical guidance on whether adjusting the cutoff values alone is sufficient or whether the models should be fully re-estimated using modern data.

3. Data and Sample Construction

3.1. Data Sources and Sample Selection

The empirical analysis is based on an archival panel dataset constructed from annual firm-level accounting data obtained from the Compustat North America Fundamentals Annual database accessed through Wharton Research Data Services (WRDS, 2026). The sample consists of standardised financial statement data for publicly listed U.S. firms covering fiscal years from 1950 up to 2025, restricted to consolidated financial statements (consol = C), the standardised data format (datafmt = STD), and industrial-format fillings (indfmt = INDL). Each observation is uniquely identified by Compustat's firm identifier (gvkey) paired with the fiscal year-end variable (fyear). This initial raw sample consists of 488,328 firm-year observations corresponding to 34,059 unique firms. Although the raw data span from 1950 onwards, the analysis sample is restricted to fiscal years 1998 to 2022 to include the three major financial crises of the twenty-first century: the Dot-com bust, the Global Financial Crisis, and the COVID-19 recession. The lower bound of 1998 is chosen to precede the Dot-com crisis, ensuring that a pre-crisis baseline is available, while the upper bound of 2022 allows to capture the full post-COVID bankruptcy wave aftermath. These form the basis of the crisis-period robustness analysis in Section 5.4.

Following Altman (1968) and Ohlson (1980), bankruptcy is defined legalistically as the filing for bankruptcy or liquidation. Consistent with prior studies that also rely on Compustat data, a firm-year observation is classified as bankrupt if Compustat's deletion-reason code (dlrsn) takes either the value 2 (bankruptcy) or the value 3 (liquidation) (Grice & Ingram, 2001; Corbae & D'Erasmo, 2021). Non-bankrupt firms comprise all remaining firm-year observations, yielding 34,059 unique non-bankrupt and 3,095 bankrupt firms. The bankruptcy reference year (failure_fyear) for the prediction models is defined as the last fiscal year for which a financial statement observation is available, i.e., failure_fyear = max(fyear), and the prediction observation is then constructed at fyear = failure_fyear - 1. Even though Compustat reports a deletion date (dldte), which denotes the calendar date on which a firm was removed from the active database, an analysis of the bankrupt sub-sample reveals that only 24% of firms have their last available financial statement year one year before to the year in which they are being deleted. This date systematically lags the last available financial statement by one to two years for the majority of firms, which could reflect both, the natural gap between a firm's fiscal year-end and the calendar date of failure, and Compustat's administrative processing time in updating deletion records, neither of which corresponds to the economic event of failure itself.

As reported in Table 1, around 60% of bankrupt firms exhibit a one- or two-year lag between their fiscal year and the deletion date. Thus, defining the reference year by max(fyear) rather than dldte decreases the missingness rate for prediction model inputs, consistent with the approach applied by Shumway (2001), and allows for a considerably larger bankrupt sample.

Table 1: Failure year definition: distribution of the lag between Compustat deletion date and last available financial statement.

Difference dldte - max(fyear)	# Bankrupt Firms	%
diff = 0 (exact match)	9	1%
diff = 1 (1 year lag)	353	24%
diff = 2 (2 year lag)	536	36%
diff > 3 (3+ year lag)	591	40%
Final Dataset	**1.489**	

3.2. Sample Construction and Splitting

The sample is subject to the following filters to obtain a clean final analysis sample:

- Deduplication at the firm-year level: where multiple observations exist for the same gvkey × fyear pair, the consolidated standardised observation with the most recent datadate is retained, dropping 13 firm-year duplicates.
- Exclusion of financial firms: observations with two-digit Standard Industrial Classification (SIC) codes in the range 6000-6999 are removed, as these hold capital structures and balance-sheet composition that are not comparable to those of industrial firms and are incompatible to the studied bankruptcy scores, dropping 134,217 firm-year observations.
- Removal of invalid denominators: observations with non-positive total assets (at ≤ 0) or non-positive total liabilities (lt ≤ 0) are dropped, since these produce mathematically undefined or economically meaningless ratios, eliminating 27,133 firm-years.
- Valid timing: keep only observations with last available fiscal-year entry pre-date the recorded deletion date (dldte), to ensure temporally consistency.

Following these filters, the sample covers the previously defined window of 1998 to 2022, removing 197,586 pre-1998 firm-year observations. Two further restrictions are then applied to construct a final common analysis sample that can run the Altman Z, Z'', and Ohlson models. First, each bankruptcy model requires valid financial data for the fiscal year immediately preceding its failure event (t−1), in order to generate a prediction for year t. Therefore, a binary indicator y = 1 is assigned to the firm-year satisfying this condition, and any bankrupt firm for which no such observation exists in the window is excluded. Second, only firm-year observations with complete data on all variables required by the Altman Z-Score, Z"-Score, and Ohlson O-Score models are retained, ensuring that all three models are evaluated on an identical set of observations. The resulting common analysis sample comprises 104,367 firm-year observations corresponding to 12,181 unique firms, of which 11,804 (96.9%) are non-bankrupt and 377 (3.1%) are firms that filed for bankruptcy at some point during the sample period. Labelled bankrupt observations (y = 1) account for 0.4% of all firm-years, reflecting the scarcity of bankruptcy in the population. Full construction details are reported in Table 2.

Table 2. Sample selection process.

Inclusion criterion	Deleted obs.	Firm-year obs.	Bankruptcies	Non-bankrupt firms
Initial raw sample from Compustat (1950–2026)	-	488.328	3.095	34.059
Single firm-year observation retained (deduplicated at gvkey × fyear	-13	488.315	3.095	34.059
Exclude financial firms (SIC 6000–6999)	-134.217	354.098	1.498	22.698
Remove invalid denominators (at ≤ 0 or lt ≤ 0)	-27.133	326.965	1.489	22.585
Validate timing (drop datadate ≥ dldte; require non-missing failure_fyear)	-23	326.942	1.489	22.585
Analysis window [1998–2022]	-197.586	129.356	526	14.147
Common scored sample for Altman Z and Ohlson O-score	-24.769	104.587	466	11.804
Require observation one year before bankruptcy (t − 1)	-220	104.367	377	11.804
Final dataset		**104.367**	**377**	**11.804**

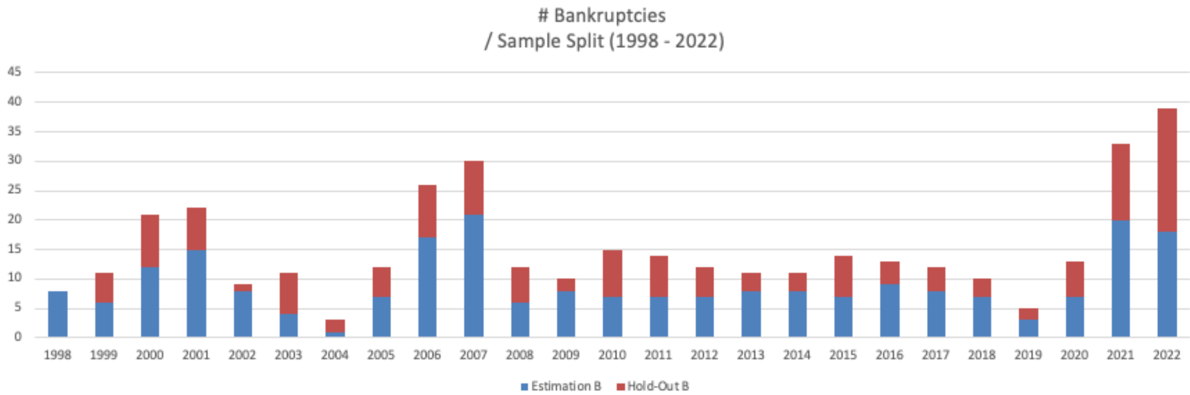
In order to replicate and re-estimate the prediction models, the sample is split into an estimation and a holdout sample. The estimation sample is used to model the new coefficients and thresholds, whilst the holdout sample is the one used to evaluate the resulting model out-of-sample. The splitting follows the methodological choices of the original Altman and Ohlson models, but both models use the same 377 bankrupt firms, divided in a 61/39 split for the estimation vs. holdout repartition, representing 229 to 148 bankrupt firms respectively. Non-bankrupt firms-year observations belonging to firms that subsequently file for bankruptcy at any point of the sample may not serve as a potential non-bankrupt control observation. Thus, no firm is used both as a control and as a bankrupt observation. Furthermore, non-bankrupt observations used to train either model are fully excluded from the holdout evaluation sample by construction. This guarantees that the observed performance differences reflect the models themselves rather than a sample overlap.

Altman's (1968) original model is estimated on a matched sample of 33 bankrupt and 33 non-bankrupt publicly traded manufacturing firms. For consistency, the estimation sample also uses a paired-sample, one-to-one, matching design, in which the non-bankrupt control firm is matched based on the same fiscal year, the same two-digit SIC code, and the closest available value of total assets (within a tolerance of plus or minus 50 per cent of the bankrupt firm's assets). In contrast, the Ohlson sample, adopts an unmatched, population-style design, which aims to approximate more closely the actual population frequency of bankruptcy than the one-to-one design used in the Altman analysis. Thus, non-bankrupt portion of the estimation sample is constructed by drawing, for each fiscal year between 1998 and 2022, a random sample of 200 non-bankrupt firms, yielding an effective bankrupt-to-non-bankrupt ratio of approximately 1:20. The same approach also yields a ratio of 1:20 for the holdout sample. The exact split can be seen in Table 3, whilst Figure 1 shows the sample repartition per year.

Table 3. Estimation and holdout sample partition for the Altman (Panel A) and Ohlson (Panel B) analyses.

	# Bankrupt	# Non-bankrupt	Ratio	Total
Panel A: Altman (1968) (paired match)				
Estimation sample	229	229	1:1	458
Holdout sample	148	2,960	1:20	3108
Panel B: Ohlson (1980) (unmatched)				
Estimation sample	229	5,000	1:20	5,229
Holdout sample	148	2,960	1:20	3,108

Figure 1. Annual bankruptcy partition by estimation and holdout samples (1998-2022, Appendix A1).



3.3. Variable Measurement and Outlier Treatment

In order to run Altman's and Ohlson's prediction models, all variables needed to compute the ratios of the linear scoring functions (1), (2), and (3), were mapped to their Compustat mnemonic conventions using standard variable definitions provided in S&P Global's Compustat Xpressfeed documentation and following previous literature of Begley et al. (1996) and Hillegeist et al. (2004) (S&P Global Market Intelligence, 2021). Table 4 provides a synthesis of the variables used.

Table 4. Variable definitions of the Altman (1968; 1995) Z-score, Z''-score (Panel A), and the Ohlson (1980) O-score (Panel B).

Variable	Definition	Compustat construction
Panel A: Altman (1968) Z-score variables		
X_1	Working capital / Total assets	(act − lct) / at; wcap when reported, fallback to act − lct
X_2	Retained earnings / Total assets	re / at
X_3	EBIT / Total assets	(pi + xint) / at; oibdp − dp when pi missing
X_4[1]	Market value of equity / Book value of total liabilities	(prcc_f × csho) / lt; for Z″, X_4_book = ceq / lt
X_5	Sales / Total assets	sale / at
Panel B: Ohlson (1980) O-score variables		
SIZE	Natural log of total assets (un-deflated)	ln(at); GNP deflator omitted (absorbed by year fixed effects)
TLTA	Total liabilities / Total assets	lt / at
WCTA	Working capital / Total assets	(act − lct) / at
CLCA	Current liabilities / Current assets	lct / act
OENEG	Indicator: 1 if total liabilities > total assets, 0 otherwise	I(lt > at)
NITA	Net income / Total assets	ni / at
FUTL	Funds from operations / Total liabilities	(pi + dp) / lt
INTWO	Indicator: 1 if net income negative for two consecutive years, 0 otherwise	$I(ni < 0 \wedge ni\_lag1 < 0)$
CHIN	Relative change in net income, scaled	$(ni_t - ni_{t-1}) / (\lvert ni_t\rvert + \lvert ni_{t-1}\rvert)$

Even after the removal of observations with non-positive total assets and total liabilities, near-zero values of these denominators can still produce extreme ratio realisations that are not

[1] Altman Z''-score calculates X_4 with book value of equity divided by book value of total liabilities and excludes X_5.

economically meaningful and were therefore replaced with missing values. Furthermore, finite but extreme ratios are addressed through winsorisation at the 1st and 99th percentiles of the distribution at the ratio level within the prediction sample, following the convention adopted in the empirical Compustat-based accounting and finance literature (Altman et al., 2018; Hillegeist et al., 2004). The binary variables OENEG and INTWO are excluded from winsorisation, as they take only the values 0 and 1 by construction.

The missingness rates in the 1998-2022 window vary across variables. For the Altman variables, rates range from under 1% for X3 and X5 to 13% for X4, driven by the lack of market equity data (mkvalt) (Appendix A2). For the Ohlson variables, SIZE, TLTA, and OENEG are fully observed due to earlier filters, while CHIN proves to have the highest rate at 7%, reflecting that there has to be two consecutive years of net income data to compute this ratio. The missingness profile is approximately balanced between bankrupt and non-bankrupt firms within each year, suggesting that the resulting selection bias is moderate. Still, the common-sample filter removes 24,769 firm-year rows, corresponding to roughly 19% of the windowed sample (Appendix A3).

3.4. Descriptive Statistics and Temporal Distribution

In order to verify that the sign and the magnitude of the cross-sectional differences observed between bankrupt and non-bankrupt firms in the modern data are broadly consistent with those documented in the original studies, as well as ensuring that the explanatory variables retain a non-trivial univariate association with the failure outcome Table 5 reports the descriptive statistics of Altman and Ohlson variables respectively.

Table 5. Descriptive statistics of Altman (Panel A) and Ohlson (Panel B) predictor variables.

Variable	Modern Estimation Sample				Original Paper			
	NB mean	B (t-2) mean	B (t-1) mean	Significant	NB mean	B (t-2) mean	B (t-1) mean	Diff. NB vs. Orig. B (t-1)[2]
Panel A – Summary statistics of Altman (1968) variables								
X_1 (WC/TA)	0.19	0.11	0.07	1.30	0.414	-	-0.061	0.25
X_2 (RE/TA)	-3.41	-5.10	-5.26	1.55	0.36	-	-0.63	-2.78
X_3 (EBIT/TA)	-0.25	-0.42	-0.43	2.11*	0.15	-	-0.32	0.07

[2] Difference = Modern NB mean minus original paper B(t-1) mean. A value near zero indicates that modern non-bankrupt firms resemble the originally distressed firms in the original study, which is the most direct evidence of structural change.

X_4 (MVE/TL)	11.21	6.80	8.39	1.09	2.48	-	0.40	10.81
X_5 (Sales/TA)	0.93	0.85	0.86	0.85	1.90	-	1.50	-0.58
Panel B – Summary statistics of Ohlson (1980) variables								
SIZE	5.11	4.63	4.55	3.73***	13.26	12.23	12.13	-7.03
TLTA	0.87	0.89	0.94	-0.90	0.49	0.72	0.91	-0.04
WCTA	-0.03	0.11	0.07	-1.54	0.31	0.16	0.04	-0.08
CLCA	1.82	1.53	1.46	1.20	0.53	0.81	1.32	0.50
OENEG	0.12	0.22	0.25	-4.58***	0.00	0.06	0.18	-0.06
NITA	-0.42	-0.55	-0.60	2.06*	0.05	-0.05	-0.21	-0.21
FUTL	-0.38	-0.90	-0.92	4.63***	0.28	-0.01	-0.12	-0.26
INTWO	0.39	0.72	0.76	-12.88***	0.04	0.18	0.39	-0.01
CHIN	-0.01	-0.10	-0.10	2.54*	0.04	0.00	-0.32	0.32

Table 6 presents the descriptive statistics for the common sample. On average, bankrupt firms are smaller with a median total asset of $126 million compared to $172 million for non-bankrupt. The mean difference of $2,355m is even larger but this is driven by a few very large non-bankrupt firms, as seen by the high standard deviation of $14,566m. Net income represents the strongest difference with a reported median of -$18m compared to +$0.3m for non-bankrupt firms. This difference explains the significance of X3 (EBIT/TA), and NITA seen in the previous Tables 5. Sales revenues report the same pattern, as the median revenue of non-bankrupt firms of $136m is more than double the one for bankrupt firms. These statistics provide evidence that there is a meaningful difference between the two groups.

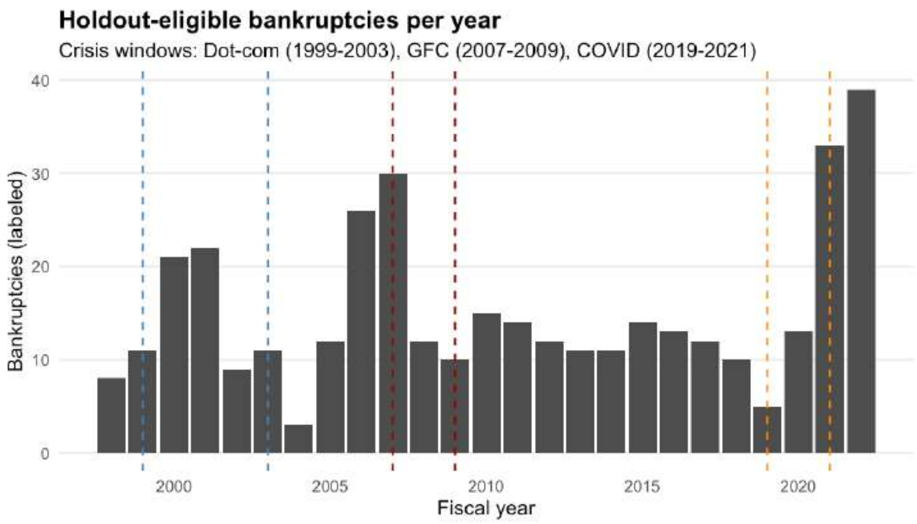
Table 6. Descriptive statistics of key firm-level financial characteristics of common sample.

	N	Mean	Median	SD	P25	P75
Panel A: Non-Bankrupt						
Total Assets ($m)	103,990	2,965	171.7	14,566	25.4	1,057
Sales Revenue ($m)	103,990	2,354	136.3	12,861	13.8	920.7
Net Income ($m)	103.990	135.1	0.3	1,435	-9.4	31.6
Panel B: Bankrupt						
Total Assets ($m)	377	609.6	126.4	2,084	26.9	412.3
Sales Revenue ($m)	377	617.7	58.3	2,842	6.2	349.3
Net Income ($m)	377	-88.1	-17.7	260.5	-80.3	-2.3

3.5. Distribution of Bankruptcy Filings over Time and across Industries

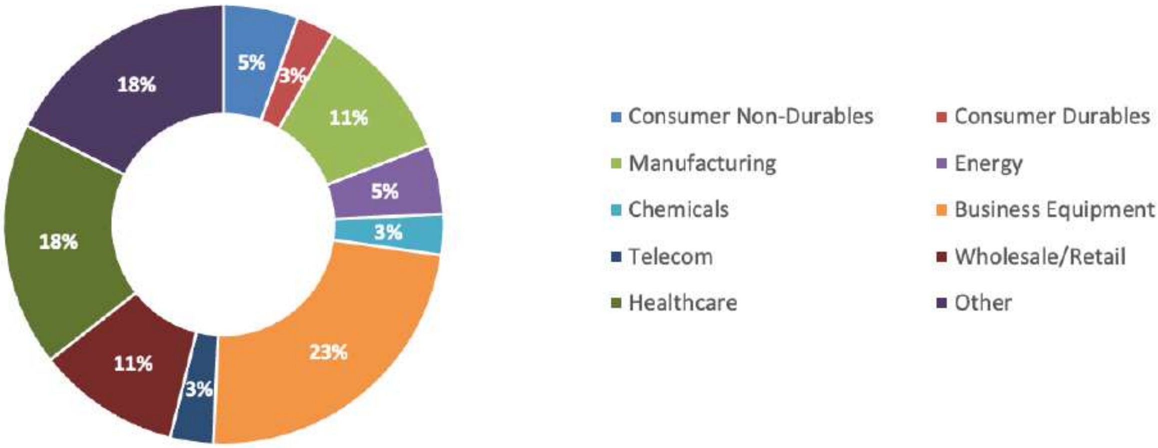
The distribution of bankruptcy filings across the analysis window can be seen in Figure 2. Thereby, the reported density proves to not be uniform, as the years 2001, 2002, 2007, 2008, 2021, and 2022 account together for around 45% of all observed bankruptcies. This finding reflects the macroeconomic impact of the Dot-com bust, the Global Financial Crisis and the Post-pandemic adjustment respectively. Additionally, the COVID-19 year of 2020 is associated with a relatively low filing count (5 filings). This is consistent with the fiscal-policy measures deployed by US authorities and the disruptions of the bankruptcy filing process during the early stages of the pandemic (Collins-Thompson et al., 2021). This results in a delay of failures into 2021 and 2022, the two years that jointly account for the highest concentration of filings in the sample with 19%. In the following analysis, the three crisis periods are defined in accordance with the US Business Cycle Expansions and Contractions, a well-established periodisation in the finance literature: (i) the Dot-com bubble (March 2001 - November 2001), (ii) the GFC (December 2007 - June 2009), and (iii) the COVID-19 crisis (February 2020 - April 2020). Moreover, a wide crisis window is defined including at least one year before each crisis, as models predict failure one year in advance, the wide window sample extends the crisis period by including additional buffer years at both ends to consider potential spillover effects. Of the 377 bankrupt firms in the initial sample, 21 (6%) represent bankrupt observations during Dot-com, 56 (15%) during GFC, and 5 (2%) during COVID.

Figure 2. Distribution of bankruptcy filings over time (1998-2022).



To analyse the sectoral impact, each firm is matched to one of the ten industry portfolios defined by Fama and French (2025), based on the standard four-digit SIC code classification. Following Fama and French's convention, SIC codes are assigned as of the end of June in fiscal year t. The distribution of bankrupt firms across the ten industry portfolios is presented in Figure 3. Business equipment, covering computers, software, and electronic equipment, constitutes the largest sector with 24,544 firm-years (23.5% of the sample) and 71 unique bankrupt firms (18.8% of all bankruptcies), followed by Healthcare and Others accounting for 18% and 17.6% of the sample, respectively.

Figure 3. Distribution of bankruptcy filings across Fama-French ten-industry groups.



4. Empirical Methodology

All computations and statistical testing are performed in RStudio version 4.4.0 (RStudio, 2026). The discriminant analysis estimator uses the *lda* function from the MASS package (Venables and Ripley, 2002), and the logit estimator uses the *glm* function with the binomial family. Data manipulation is performed using *dplyr* and *tidyr*, visualisation using *ggplot2*, and output validation using *testthat*.

The metric used for evaluating the discriminatory validity of each model is classification error rates, consistent with the evaluation framework applied in the original studies (Altman, 1968; Ohlson, 1980). Three metrics are reported: (i) Type I error, representing the percentage of bankrupt firms that are incorrectly classified as non-bankrupt (missed bankruptcies) (= FN / (TP + FN)); (ii) Type II error, representing the percentage of non-bankrupt firms incorrectly classified as bankrupt (false alarms) (= FP / (TN + FP)); (iii) Combined error representing the unweighted sum of Type I and Type II error rates, which assigns equal weight to both error types regardless of class imbalance (= (Type I + Type II) / 2). Additionally, Sensitivity (= 1 - Type I) and specificity (= 1 - Type II) are reported to facilitate comparison with the broader distress prediction literature. The combined error metric is deliberately unweighted, as an observation-weighted metric would be dominated by the non-bankrupt majority and mechanically minimise by classifying all firms as non-bankrupt. All performance metrics are computed on the common holdout sample of 148 bankrupt and 2,960 non-bankrupt firm-year observations, to ensure unbiased comparability.

In addition, the stability of classification rankings across model specifications is assessed using a bootstrap robustness analysis. The holdout sample is resampled with replacement across 100 iterations, and Type I, Type II, and combined error rates are recomputed for each specification in each iteration. The resulting distributions of classification errors are used to test the robustness of model performance rankings to variation in holdout sample composition.

4.1. Multiple Discriminant Analysis (Altman Z and Z'')

Altman's (1968) model applies Multiple Discriminant Analysis (MDA), introduced by Fisher (1936), to predict bankruptcy. This approach derives a linear combination of financial ratios that maximally separates bankrupt from non-bankrupt firms into a single discriminant score (the Z-score). A firm-year is classified as bankrupt if its score falls below a published cutoff. To evaluate the original model on the modern holdout sample, the published coefficient vectors are applied to each holdout firm-year observation (Section 2.1.). For the Z-score, the primary cutoff of 2.675 and the lower distress cutoff of 1.81 are tested, all firm-year observations below these cutoffs are classified as bankrupt. For the Z''-score (Altman, 1983), the equivalent cutoffs of 2.60 and 1.10 are applied.

To ensure statistical replication and comparability, both Altman models are re-estimated by refitting via the same linear combination of a vector of predictor variables x that maximises the ratio of the between-group variance to the within-group variance on the matched estimation sample. As described in Section 3, all 229 bankrupt estimation firms are matched at the 2-digit SIC level with 229 non-bankrupt firms using equal group priors consistent with the balanced sample design. Matching Altman's convention, scores are sign-normalised so that lower values denote more distress. The optimal cutoff C* is determined on the estimation sample as described in Section 4.4.

4.2. Binary Logistic Regression (Ohlson O-score)

Ohlson's (1980) model applies a conditional logit specification, which models the probability of bankruptcy as a function of nine financial predictors. Unlike MDA, a logistic regression does not require multivariate normality or equal variance-covariance matrices across groups, making it theoretically better suited to the distributional properties of financial ratios (Press and Wilson, 1978). Parameters are estimated by maximum likelihood, providing a percentage result, instead of a score. The published coefficient vector is applied to each holdout observation to compute a predicted bankruptcy probability (see Section 2.1.). In Ohlson's model, a firm-year is classified as bankrupt if this probability exceeds the published cutoff of 0.038. Furthermore, a cutoff sensitivity analysis across the full probability range to provide insights on classification performance based on the cutoff choices.

For re-estimation, the model is refitted by maximum likelihood logistic regression on the unmatched estimation sample of 229 bankrupt and 5,000 non-bankrupt firm-years described in Section 3. Winsorisation is applied to the seven continuous predictors using bounds derived from the estimation sample only, excluding binary variables OENEG and INTWO. The optimal cutoff C* is determined on the estimation sample as described in Section 4.4.

4.3. Cutoff Optimisation (C*)

For each re-estimated model, an optimal classification cutoff C* is identified on the estimation sample by performing a grid search over all candidate cutoff values and selecting the one value that minimises the unweighted sum of Type I and Type II error rates, assigning equal importance to missed bankruptcies and false alarms and preventing that all firms are classified as non-bankrupt. Once determined on the estimation sample, C* is held fixed and applied unchanged to the holdout, ensuring that the cutoff does not incorporate any holdout information.

5. Empirical Results

5.1. Original Model Performance on Modern Data

Applying both of the original Altman models to the modern holdout sample discloses that their performance decreases significantly (Table 7). At its primary cutoff of 2.675, the Altman Z-score achieves a Type I error of 23.0%, meaning that it correctly identifies 77% of bankruptcies, but at the cost of a high false-alarm rate (Type II: 52.0%). Thus, the original model flags more than half of all healthy firms as distressed, resulting in a combined error of 37.5%. When using the lower distress cutoff of 1.81, the trade-off shifts. Type I error rises to 31.1% while Type II falls to 39.7%, yielding a combined error of 35.4%. Although this is a means an improvement of 2%, the model still misses nearly one in three bankruptcies, and the combined error remains almost three times as high as the 12.5% reported by Altman (1968) on his original holdout sample. One likely explanation is the structural gap between the two estimation windows. The original coefficients were calibrated using U.S. manufacturing firms from 1946 to 1965, a period characterised by more asset-intensive corporate structures and substantially different accounting and financing regimes. Applying those historical coefficients to a modern 1998-2022 universe, which includes a larger sample of firm types, previous literature documents that financially healthy firms naturally exhibit low Z-scores because their financial profiles have changed, not because they are in distress (Begley et al., 1996). This shifts the entire score distribution downward, increasing Type II errors irrespective of the cutoff applied.

Table 7. Out-of-sample performance of the original Altman Z- and Z''-scores on the modern holdout sample (in %).

Model	Type I	Type II	Combined	Sensitivity	Specificity
Panel A – Altman Z-score (1968)					
Original Altman (Z = 2.675 cutoff)	23.0	52.0	37.5	77.0	48.0
Original Altman (Z = 1.81 cutoff)	31.1	39.7	35.4	68.9	60.3
Panel B – Altman Z''-score (1983)					
Altman (Z'' = 2.60 cutoff)	20.9	54.6	37.8	79.1	45.4
Altman (Z'' = 1.10 cutoff)	30.4	40.4	35.4	69.6	59.6

The Altman Z"-score (Altman, 1983) was designed to broaden the model's applicability beyond manufacturing by excluding the asset turnover ratio (X5) and replacing market value equity for book equity in the leverage ratio (X4), making it better suited to a diverse sample. Although previous literature suggests that these changes improve accuracy on modern datasets, it unexpectedly does not perform better than the original Z-score here (Altman, 2018). At the

2.60 cutoff, the combined error is 37.8% (Type I: 20.9%; Type II: 54.6%) and when lowering the cutoff to 1.10, the combined error is reduced to 35.4% (Type I: 30.4%; Type II: 40.4%), an identical pattern to that of the Z-score. Hence, the Z"-score's failure to improve on the original model suggests that the performance limitation is not the inclusion of asset turnover ratio nor the usage of equity value, but the underlying problem of coefficient calibration on an unrepresentative historical sample. Both models illustrate the same fundamental trade-off pattern between Type I and Type II error, which reappear throughout other analyses in this thesis. At the higher cutoffs (2.675 or 2.60), the models capture most bankruptcies, with Type I errors around 21-23%, but classify more than half of all healthy firms as distressed, yielding Type II errors above 50%. Using a lower cutoff (1.81 or 1.10) maximises sensitivity at the cost of specificity, with false alarms falling meaningfully but still, nearly one in three bankruptcies is missed. Neither calibration produces a result that allows proper discrimination between bankrupt and non-bankrupt firms on modern data and is suitable as standalone screening tool.

Table 8. Out-of-sample performance of the original Ohlson (1980) O-score on the holdout sample (in %).

Model	Type I	Type II	Combined	Sensitivity	Specificity
Original Ohlson (cutoff = 0.038)	2.7	84.3	43.5	97.3	15.7

The Ohlson O-score at its original cutoff of 0.038 presents the most extreme case. Type I error is at 2.7%, meaning 97.3% of actual bankruptcies are correctly flagged, but this comes at the cost of a higher Type II error (84.3%), meaning that the model classifies the majority of healthy firms as distressed (Table 8). A cutoff grid (Appendix Table B2) illustrates that this failure is almost entirely attributable to the choice of cutoff rather than the underlying model. When raising the cutoff to 0.6, the original Ohlson coefficients achieve a combined error of 34.6%, with a far more balanced Type I of 29.7% and Type II of 39.4%. In other words, the cutoff appears miscalibrated for the 1998-2022 time period. This temporal instability of optimal logistic cutoffs is a known limitation of static model applications and has been documented in prior literature (Shumway, 2001).

5.2. Re-estimation of Bankruptcy Models

After re-estimating the Altman Z-score, results indicate that the new coefficients diverge from Altman's (1968) original discriminant function (Table 9, Panel - A). While the original five variables' weights are all positive, with X1 = 1.20, X2 = 1.40, X3 = 3.30, X4 = 0.60, X5 = 1.00, the re-estimated function, after sign normalisation, produces a negative weight for X2 and shows an overall reduction across all five variables. The most notable difference is the sign reversal on X2 (retained earnings/total assets). Historically, higher retained earnings relative to assets indicated accumulated profitability, financial resilience, and a lower probability of bankruptcy. In the modern sample, this discriminatory signal has weakened, as non-bankrupt firms also report negative X2 values (mean X2 = -3.40 for non-bankrupt firms in the estimation sample).

This shift can be explained by a fundamental change in the composition of the Compustat universe since the original estimation period. An analysis of the share of technology and intangible-intensive firms in Compustat by five-year interval reveals that this share grew from 3.5% in 1950-1954 to 21.8% by 2000-2004 (Appendix B5). As a consequence, the modern Compustat sample is heavily populated by asset-light firms that tend to reinvest their capital into growth before reaching profitability, resulting in oftentimes having negative retained earnings. Here, the trade-off is faster growth and scalability of this firms, not a signal of distress. This is consistent with the findings of Fama and French (2004) in their study of newly listed US firms. They documented that lower equity financing costs allowed weaker firms and firms with more distant expected payoffs, such as technological or growth-stage companies, to access public markets from the 1980s onwards, making newly listed firms progressively more left-skewed in profitability. One concrete example is Amazon. After it went public in 1997, it reported negative retained earnings for its first eleven years (from 1997 - 2008), at the same time, it continued to expand rapidly and maintained access to external financing (Appendix B3). This composition shift is further reflected in the industry distribution of the common sample used in this thesis, where Business Equipment, the Fama-French category that covers computers, software, and electronic equipment, represents the largest single sector, accounting for 23.5% of all firm-years and 18.8% of the bankrupt subsample (Figure 3).

Re-estimated Altman Z"-score coefficients show a similar pattern, with X2 again reversing sign (Table 9 - Panel B). Overall, Z’’ coefficients decrease even more heavily than those of the Z-score with an average reduction of 93% across all variables. The strongest remaining discriminator is X3 (EBIT/Assets), and the weakest is X4 (Book Common Equity/ Liabilities), which collapses almost to zero (X4 = 0.02). This is consistent with the univariate evidence in Table 5, where X3 is the only Altman variable that achieves significance. Unlike the Ohlson probability cutoff, the re-estimated Altman cutoffs are not directly comparable to the published cutoffs, as re-estimation produces scores on a different discriminant score scale.

Table 9. Coefficient comparison - Altman (1968) Z-score and Altman (1983) Z''-score.

	X5	X2	X3	X4[3]	X5[3]	Cutoff
Panel A - Altman Z						
Original Altman Z	1.20	1.40	3.30	0.60	1.00	2.675
Re-estimated Altman	0.72	-0.03	1.00	0.01	0.37	0.661
Difference (re-est. − orig.)	-0.48	-1.43	-2.30	-0.59	- 0.63	-2.014
Panel B - Altman Z''						
Original Altman Z''	6.56	3.26	6.72	1.05	-	2.600
Re-estimated Altman	0.67	-0.03	1.18	0.02	-	0.107
Difference (re-est. − orig.)	-5.89	-3.29	-5.54	-1.03	-	-2.493

The re-estimated Ohlson O-score reveals even larger structural changes (Table 10). The largest coefficient change occurs for TLTA, as it decreases from 6.03 to 0.210, representing a 97% reduction on modern data. An analysis of the evolution of median TLTA across five-year intervals from 1950 to 2024 for non-bankrupt and bankrupt firms in the Compustat universe reveals that the median leverage ratio of non-bankrupt firms rose from an average of 0.36 in the 1950s-1965s to approximately 0.55 in the 2000s-2020s, confirming a 47% structural increase over this period (Appendix B4). This distributional shift is further evidenced in Table 5, which shows that the average non-bankrupt firm today carries a mean TLTA of 0.87, almost double the average leverage held by a bankrupt firm in Ohlson's original non-bankrupt sample (TLTA = 0.49). This convergence explains the variable's reduced ability to separate the two groups. Consequently, the modern non-bankrupt median leverage of 0.55 is nearly identical to the median leverage of bankrupt firms during Ohlson's (1980) original estimation window of 1970-1974 (0.56). Thus, leverage operates at levels that would have been characteristic of distressed firms which supports that leverage has significantly reduced its discriminatory validity in modern period. This may be explained by the low cost of debt, as US 10-year Treasury interest rates declined continuously from 10% in 1980 to under 2% by 2022 and the growth of debt covenant renegotiation as an alternative to formal bankruptcy (FRED, 2025; Roberts et al., 2009). Both reduce the risk historically associated with higher leverage and make it more attractive to carry higher debt levels, explaining the rise in observed leverage levels. With an almost sixfold increase from 0.285 to 1.680, INTWO, which represents two consecutive years of operating losses, is the strongest predictor of bankruptcy in the re-

[3] For the Z''-score, X4 is based on the book value of equity instead of market value, while X5 is omitted from the model

estimated Ohlson model, consistent with its univariate dominance in Table 5 ($t = -12.88$, $p < 0.001$). Whereas WCTA, CLCA, and NITA reverse signs, the correlation matrix reveals near-perfect collinearity between TLTA and WCTA ($r = -0.967$) and strong collinearity between TLTA and CLCA ($r = 0.806$), meaning their individual coefficient signs become unreliable (Appendix A4).

Table 10. Coefficient comparison - Ohlson O-score (1980).

	SIZE	TLTA	WCTA	CLCA	OENEG	NITA	FUTL	INTWO	CHIN	Cutoff
Original Ohlson O	-0.41	6.03	-1.43	0.08	-1.72	-2.37	- 1.83	0.29	-0.52	0.038
Re-estimated Ohlson	0.05	0.21	0.26	- 0.01	0.74	0.07	-0.00	1.68	-0.35	0.051
Difference (re-est. - orig.)	0.46	- 5.82	1.69	-0.09	2.46	2.44	1.83	1.39	0.17	0.013

Before assessing out-of-sample classification performance, the overall goodness of fit of the re-estimated Ohlson model is examined using the likelihood ratio index (LRI), defined as LRI = 1 − (log-likelihood at convergence / log-likelihood at zero), consistent with Ohlson's (1980) original reporting convention. The re-estimated model achieves an LRI of 76.5% on the modern estimation sample, compared to Ohlson's (1980) original LRI of 84.0%, which represents a 7.5 percentage point decline that confirms that six out of the nine predictor variables retain meaningful explanatory power when coefficients are updated to the modern period. By contrast, applying Ohlson's (1980) published coefficients directly to the modern estimation sample yields an LRI of -322.8%, meaning the original coefficient vector fits modern data worse than a model that assigns a 50% bankruptcy probability to every firm. A formal likelihood ratio test confirms that re-estimation is statistically necessary (LR = 28,942; $p < 0.001$), ruling out the interpretation that the original coefficients require only minor adjustment.

A potential concern is whether these observed coefficient changes reflect genuine structural shifts in the relationship between financial ratios or merely changes in the distribution of the predictor variables. Distributional shifts as documented in Table 5 could mechanically reduce coefficient magnitudes without any change in the underlying relationship. Thus, to test if this holds, standardised coefficients are computed (= raw coefficient * standard deviation of the variable in the non-bankrupt group), allowing a direct comparison across time (Appendix B7). Even though the coefficient change for CHIN is largely attributable to distributional scaling (standardised effect change: -17%), the analysis shows that the changes in INTWO (+1,310%), OENEG (+310%), and NITA (+156%) represent genuine structural change that cannot be explained by shifts in variable distributions alone. For TLTA, distributional change accounts for approximately two-thirds of the raw coefficient decline. Nevertheless, the standardised effect still falls by 68%, confirming a meaningful structural weakening of leverage as a bankruptcy predictor beyond what distributional scaling would imply. SIZE becomes positive (0.053 versus original -0.407). However, SIZE carries no interpretable meaning as it experiences multicollinearity and non-bankrupt firms still prove to have higher mean SIZE (B = 4.55 vs. NB = 5.11, $t = 3.73$, $p < 0.001$), consistent with Ohlson's original finding. The optimal O-score cutoff C* = 0.051, modestly higher than the original 0.038.

5.3. Model Performance Comparison

Despite re-estimating the coefficients on modern data, the re-estimated Altman Z-score underperforms relative to the original Z-score, as the combined error increases from 35.4% (Z orig. 1.81) to 43.1% (Z re-est, C*). While it is a counterintuitive result, this is consistent with Begley et al. (1996) findings. They found that re-estimating Altman's model on out-of-sample data from the 1980s does not reliably improve the combined error rate, particularly when violating the multivariate normality assumption. This emphasises that the performance degradation reflects a deeper limitation of the variable set rather than a simple miscalibration of coefficients. Even when estimating on the same five Altman variables with an auxiliary logistic regression, a methodology that does not require normality or equal covariance matrices, the same results appear. None of the five variables achieve statistical significance (all $p > 0.10$, Appendix A5), confirming that the Altman variable set has lost discriminatory power on modern data. In contrast, the re-estimated Z'' model performs considerably better than its original model, with a combined error rate of 35.2% (Type I: 34.5%; Type II: 35.9%; C* = 0.107), and proves to perform best among the six Altman specifications.

Overall, the best performance is achieved by the re-estimated Ohlson O-score at C* = 0.051 with a combined error of 30.4%. This represents a meaningful improvement over the original O-score at 0.038 (combined: 43.5%) and over the Altman models at their best specifications. The modest upward shift in cutoff from 0.038 to 0.051 reduces the high Type II error of the original cutoff by around 50 percentage points (from 84.26% to 34.43%).

Table 11. Comprehensive model performance comparison on the modern holdout sample (in %).

	Cutoff	Type I	Type II	Combined	Sensitivity	Specificity
Ohlson re-estimated (C)*	0.051	26.35	34.43	30.39	73.65	65.57
Z" re-estimated (C)*	0.107	34.46	35.88	35.17	65.54	64.12
Z original (1.81)	1.810	31.08	39.66	35.37	68.92	60.34
Z" original (1.10)	1.100	30.41	40.44	35.42	69.59	59.56
Z original (2.675)	2.675	22.97	52.03	37.50	77.03	47.97
Z" original (2.60)	2.600	20.95	54.56	37.75	79.05	45.44
Z re-estimated (C*)	0.661	29.05	57.06	43.06	70.95	42.94
Ohlson original (0.038)	0.038	2.70	84.26	43.48	97.30	15.74

Table 11 summarises all eight model specifications, confirming that the three best combined performances are obtained by the re-estimated O-score, the re-estimated Z" at C* = 0.107, and the original Z at 1.81, achieving 35.37%. While, in absolute terms, a combined error rate around 30% remains substantially worse than the error rates reported in the original studies, direct comparison is limited by differences in sample construction. Moreover, as shown in the descriptive statistics in Table 5, only 1 out of 5 Altman variables and 5 out of 9 Ohlson variables are statistically significant, which indicates that the variables genuinely struggle to discriminate between the groups on modern data.

To test the robustness of these results, the same analysis was repeated using raw ratios and discloses that winsorisation has minimal impact across all specifications. Raw and winsorised variants differ less than 2 percentage points in all cases (Appendix B1). Furthermore, a bootstrap robustness analysis using 100 resampling iterations (Table 12) confirms the stability of the rankings. The re-estimated O-score at C* still achieves the best mean combined error of 31.5% with the tightest bootstrap interval of the best performing specifications[4] (IQR: [27.8%, 34.8%]) with a range of 1.4 percentage points. It indicates that it is the best performing and the most stable specification across different out-of-sample bootstrap iterations. On the other hand, the re-estimated Z" at C* shows a wider sampling interval (IQR: [35.1%, 39.6%]), suggesting greater sensitivity to holdout composition. The original Z"-score at C = 1.10 achieves the second-best bootstrap mean of 33.3% (IQR: [32.4%, 34.2%]), followed closely by the original

[4] The original O-score at C = 0.038 technically achieves the tightest interval of 0.8%, but with a mean combined error of 43.9%, given its poor predictive performance it is excluded from the stability comparison.

Z-score at C = 1.81 with a mean of 33.6% (IQR: [32.5%, 34.3%]). Overall, this analysis confirms that the re-estimated Ohlson model dominates across both performance and stability.

Table 12. Bootstrap robustness analysis: distribution of combined error rates across 100 resampling iterations (in %) (extract of Z-score at C* = 1.81 and Ohlson re-estimated results).[5]

Combined Error	Mean	Median	Q25	Q75	Min	Max
Z original (1.81)						
Type I	27.0	27.0	25.7	28.4	18.9	33.1
Type II	40.1	40.0	39.4	40.6	37.5	42.2
Combined	33.6	33.5	32.5	34.3	29.3	37.1
Ohlson re-estimated (C)*						
Type I	23.3	23.6	16.2	28.4	9.5	39.2
Type II	39.8	37.2	35.1	46.0	25.5	51.0
Combined	31.5	31.5	30.9	32.3	27.8	34.8

5.4. Crisis Period Performance

A central contribution of this thesis is the evaluation of the models' performance during periods of macroeconomic crisis relative to non-crisis periods, using well-established periodisation in the finance literature. Firms whose bankruptcies fall within a crisis event are classified as crisis-period observations, such that the prediction-year observation (fyear = failure_fyear - 1) falls one year prior to the recession: fyear = 2000 for the Dot-com bust, fyear = 2006-2008 for the GFC, and fyear = 2019 for the COVID-19 pandemic. All remaining observations form the non-crisis group. The objective is to assess whether the bankruptcy prediction models, applied one year before the crisis, can correctly identify firms that will subsequently fail during that crisis. Table 13 reports the sample split per crisis event used for the following analysis.

Table 13. Crisis period sample composition: narrow and wide crisis windows definitions.

# Bankrupt	Narrow window	Wide window
Dot-com	9	22
Great Financial Crisis	24	31
COVID-19	2	21
Total Crisis	**35**	**74**
Non-Crisis	113	74

[5] Full table can be seen in Appendix B6.

To account for the aftermath of each crisis, a second, wider window extends the NBER recession dates. This is particularly important for the COVID-19 crisis, as the narrow window, containing only 2 bankruptcies, would not be interpretable.

Table 14. Best combined error performance by crisis period: narrow and wide window results (in %) (Extended version: Appendix B9).

Period	#N Bankrupt	O Re-est. C*	Z''-Re-est. C*	Z – Original (1.81)
Panel A – Narrow Window				
Dot-com[6]	9	25.0	37.0	44.6
Great Financial Crisis	24	29.1	40.5	43.9
COVID[6]	2	17.4	21.6	21.6
Panel B – Wide Window				
Dot-com	22	26.5	34.9	38.1
Great Financial Crisis	31	31.0	38.3	43.0
COVID	21	35.1	31.9	32.9
Normal	**113**	**31.2**	**34.0**	**33.0**

Looking at Table 14, one of the most notable findings is that, under the best-performing specifications, neither the GFC nor the Dot-com periods provide evidence that these periods are materially more difficult to predict than normal times. The COVID-19 period represents the only exception. Under the re-estimated O-score, GFC achieves a combined error of 29.1% in the narrow window and 31.0% in the wide window, both below the normal-period result of 31.2%. The Dot-com wide window similarly shows 26.5%, confirming the pattern. While its narrow window with N = 9, should be treated with caution, the wide- and narrow-window results difference of only 1.5 percentage points (26.5% versus 25.0% for the re-estimated O-score), confirms that the Dot-com finding is robust even when the sample size more than doubles. The only exception is COVID-19, where the re-estimated O-score reaches 35.1%, which is above the normal-period of 31.2%. Here, the best specification is realised by the re-estimated Z''-score with a combined error of 31.9%. This discloses the only period in which Altman specifications attain lower combined errors than the re-estimated O-score, suggesting that the Altman variable set captures COVID-19 distress signals more effectively than the Ohlson model. Nonetheless, particularly during GFC, the Altman models perform considerably

[6] COVID-19 and Dot-com narrow window contain only N = 2 and N = 9, respectively, making their narrow-window sample size statistically unreliable.

weaker than in normal times, as the original Z model reaches more than five percentage points difference (38.1% versus 33.0%).

These findings are counterintuitive. Bankruptcy prediction models evaluate the financial health of individual firms. Consequently, one would expect them to perform worse during systemic stress periods, when failures may be driven by external macroeconomic shocks rather than by the gradual deterioration of firm-specific financial ratios. Remarkably, the results do not indicate a material decline in predictive performance during the Dot-com and GFC periods, which can be explained by the following reasons. First, financial profiles of bankrupt firms appear broadly similar, regardless of whether the failure occurs during a crisis or during normal times (Appendix B8). The year-before-bankruptcy characteristics observed in the holdout sample suggest that their financial distress was already building up prior to the crisis event itself, implying that the fundamental warning signals used in the prediction models follow a relatively consistent trajectory independently of the broader macroeconomic environment. This trajectory is consistent for both crisis-period and normal-period failures, suggesting that crisis conditions may merely accelerate failure for firms already on a distress path.

Second, financial institutions, which were among the firms most severely affected by the GFC, were excluded from the common sample by construction (i.e. SIC filter in Table 2). During the GFC, the systemic credit market freeze particularly caused financial firms, that relied on short-term debt and were unable to refinance their obligations, to fail. Subsequently, the GFC holdout subsample primarily consists of non-financial firms whose financial distress was likely already exhibited prior to the crisis. Thus, these firms may not fully represent the systemic bankruptcies typically associated with crisis times, but rather failures that would have occurred independently of the crisis itself.

Third, the Compustat universe used in this analysis is restricted to publicly listed firms. These are typically larger and have easier access to external financing compared to privately held companies. Therefore, public firms can absorb temporary crisis more easily as opposed to smaller firms that more often depend on bank lending. Their vulnerability might affect them during crisis, such as credit-supply shocks. This effect was particularly striking during COVID-19, as many smaller firms such as restaurants or small retails, which were not fundamentally insolvent beforehand, failed. Consequently, systemic crisis would rather accelerate or disclose a public firm’s financial weaknesses rather than causing bankruptcy, when no previous financial distress was detected. This supports the continued usage of accounting-based bankruptcy models, as their predictive ability holds. Lastly, the subsample for the narrow-window defined crisis periods comprises only a small number of observation (2 to 24), which means that the results for individual firm carry a disproportionate weight on the error rates and should be interpreted with caution.

5.5. Industry-Level Performance

As Altman originally used a sample of manufacturing firms only to predict his model, Ohlson developed his model on a broader industrial sample, only excluding utilities, transportation companies, and financial services companies as these industries are structurally different (Altman 1968; Ohlson, 1980). In order to assess whether Altman's model generalises beyond manufacturing to the diversified modern firm population, and which model performs best for which type of industry, the following section tests whether predictive performance varies across industries and to what extent. The holdout sample is reclassified according to the Fama-French ten-industry scheme. Among the 148 holdout bankruptcies, Healthcare is the most represented sector with 24% (N = 35), followed by Business Equipment, Other, and Wholesale and Retail, representing 19%, 18%, and 16% respectively (Figure 4). An important limitation is that six of the ten industry groups contain less than 10 bankrupt observations[7].

Figure 4. Industry distribution of holdout bankrupt firms by Fama-French classification.

Table 15 presents the best achievable combined error per industry across all eight bankruptcy prediction model combinations. The results show that seven of the ten industries achieve the lowest combined error when using the re-estimated Ohlson model, confirming its general superiority documented in the aggregate results in Section 5.3. The three exceptions are Manufacturing (original Z''-score at C = 1.10), and Healthcare and Energy, both with the original Z-score (C = 1.81). Remarkably, the three sectors in which Altman specifications perform better, are manufacturing and capital-intensive industries, which is exactly the environments in which Altman was originally calibrated on.

Table 15. Model performance by industry: Fama-French ten-industry classification (in %) (Extended version: Appendix B10).

Industry	#N bankrupts	Best combined error	Type I error	Type II error	Best prediction model
Consumer Durables[7]	5	12.5	0.0	25.0	O re-est C*
Telecom[7]	4	24.5	0.0	49.1	O re-est C*
Business Equipment	28	25.8	10.7	40.9	O re-est C*
Manufacturing	8	26.0	25.0	27.0	Z" origin. C = 1.10
Healthcare	35	28.4	11.4	45.3	Z origin. C = 1.181
Energy	9	28.8	0.0	57.7	Z origin. C = 1.181
Other	26	29.7	34.6	24.9	O re-est C*
Wholesale/Retail	24	32.8	45.8	19.8	O re-est C*
Consumer Non-Durables	7	36.4	57.1	15.6	O re-est C*
Chemicals[7]	2	39.7	50.0	29.5	O re-est C*

The analysis shows that the Type I to Type II trade-off depends widely on the industries. The highest Type II errors are observed in the Energy and Healthcare sectors, with 57.7% and 45.3% respectively, meaning both models flag a very large share of healthy firms as distressed. For Energy, the original Z-score at C = 1.81 catches all bankruptcies (Type I = 0%) but at the cost of flagging more than half of healthy energy firms, which is consistent with the finding that energy firms carry volatile financial ratios that can resemble distress signals even during periods of financial health. In contrary, the highest Type I errors are found in Consumer Non-Durables and Wholesale and Retail, with 57.1% and 45.8% respectively, indicating that these models miss a large share of actual bankruptcies in these sectors.

Looking at Manufacturing, the industry used exclusively to construct the original Altman model, the best combined error achieved is 26% under the original Altman Z''-score at its 1.10 cutoff. With a Type I of 25% and Type II of 27%, this is also the most balanced Type I/II profile of all and represents the second lowest result across the sectors with large enough sample sizes after Business Equipment. While Altman (1968) reported a better holdout combined error of 12.5% on a manufacturing-only sample, direct comparison is limited by the

[7] Results for Chemicals, Telecom, and Consumer Durables are reported as indicative but are not reliably interpretable given the small sample sizes.

differences in time period, firm characteristics, and sample composition. between the two studies. This finding supports also the expectation that Altman models perform best within their original calibration domain. The strongest overall performance is achieved by Ohlson's re-estimated model on Business Equipment with 25.8% combined error. Among industries with interpretable sample sizes, Wholesale and Retail (N = 24, combined error 32.8%) show the weakest performance. These sectors would likely benefit from specialised modelling approaches rather than generic ratio-based classifiers.

5.5.1. Tech Sector Analysis

The emergence of technology-intensive, asset-light business models since the late 1990s with the rise in technological innovations, such as the internet, represents one of the main structurally different industries compared to the capital-intensive manufacturing firms used for original model calibration. Technology firms, classified as Business Equipment under Fama-French (FF10 = 6), correspond broadly to SIC codes 3570-3699 and 7370-7379. These typically carry large, accumulated deficits, negative net income, and minimal tangible assets, as a consequence of their business models and their growth-oriented investment strategies. Moreover, they represent the sector with the highest number of bankruptcies in the sample. This raises the question of whether traditional financial ratios retain a discriminatory validity within this sector.

To understand whether bankrupt tech firms are a structurally distinct failure type and are financially distinguishable from non-tech bankruptcies, a Welch t-test is conducted for each variable separately on the subset of bankrupt holdout firms. Thus, comparing the 28 tech bankruptcies against the remaining 120 non-tech bankruptcies (results reported in Appendix Table B11). Of the fourteen variables tested, only INTWO is statistically significant and different ($t = 2.86$, $p = 0.006$), as 89.3% of bankrupt tech firms experienced two consecutive years of net losses prior to failure, compared to 68.3% among non-tech bankruptcies. This is further supported descriptively by the mean net income to total assets of -3.07 among bankrupt tech firms, compared to -0.47 among non-tech bankruptcies, suggesting a cash-burn trajectory rather than a leverage shock or margin collapse as the dominant failure mechanism.

Furthermore, to test whether bankrupt and non-bankrupt tech firms can be separated using the traditional variables, both groups are compared within the holdout sample. Results reveal that only three of the fifteen model variables significantly discriminate between the two groups (Table 16), compared to four in the full holdout sample, meaning that the tech sector is harder to discriminate within than the full sample using these traditional variables. The strongest discriminator is X4, market value of equity to total liabilities with a p-value of < 0.001, where bankrupt tech firms have a mean of 4.16 versus 11.39 for healthy firms. This reflects the tendency of healthy technology companies to command high market valuations relative to their debt, whilst distressed firms see their market capitalisation collapse ahead of failure. As X4 is a market-based rather than purely accounting-based signal, this finding also suggests that forward-looking market information is more informative for tech bankruptcy prediction, aligning with previous findings (Hillegeist et al., 2004). INTWO ($p < 0.001$) and OENEG ($p = 0.035$) are the remaining significant variables.

Table 16. Technology firms' variable comparison for bankrupt versus non-bankrupt firms.

Variable	Bankrupt mean	Non-bankrupt mean	Significance
X4 (MVE/TL)	4.16	11.39	***
INTWO	89.3	43.3	***
OENEG	32.1	12.0	*
Everything else	-	-	Not significant

The two variables that carry significant discriminatory power in the full holdout but lose their signal within the technology sector are FUTL (funds from operations coverage) and CHIN (change in net income). This is likely because healthy technology firms more often show poor cash flow metrics and declining earnings because they are investing aggressively in growth. This supports the argument that tech-sector bankruptcy prediction could benefit from replacing traditional variables with market-based or tech-specific variables, such as R&D intensity, Cash burn rate, Revenue growth rate, Intangible assets/ Total Assets, Gross margin, Operating leverage/ Fixed cost ratio, or Equity issuance/ Financing dependence, and testing whether this improves the model.

6. Conclusion

This study examines whether the traditional bankruptcy prediction models Altman's Z-score (1968), Altman's Z"-score (1983), and Ohlson's O-score (1980) retain their predictive validity on modern US data (1998-2022), whether re-estimating their parameters improves their accuracy, and how performance varies across crisis periods and industry groups.

The results show that all three original bankruptcy prediction models report significant performance degradation on the modern holdout sample, with combined errors of original specifications ranging from 35.4% (Z-score at C = 1.81) to 43.5% (O-score at C = 0.038). Re-estimating coefficients and recalibrating cutoffs partially recover predictive power, such that the best-performing specifications retain meaningful discriminatory ability. The best-performing specification is achieved by the re-estimated Ohlson O-score at C* = 0.051 producing a combined error of 30.4%, a result confirmed by a bootstrap robustness analysis across 100 iterations which also identifies the re-estimated Ohlson O-score as the most stable specification with an IQR of [30.9%, 32.3%]. The re-estimated Altman Z"-score at C* = 0.107 follows with a combined error of 35.2%. No specification realises a combined error below 30% implying that approximately one in three classifications will be incorrect. These results broadly align with Begley et al. (1996), who document a similar advantage for the Ohlson model when applied to out-of-sample data from the 1980s. The traditional bankruptcy models should therefore not be relied upon as standalone tools for identifying firms at risk of bankruptcy.

A counterintuitive finding of this study is that re-estimating the Altman Z-score on modern training data worsens, rather than improves, out-of-sample classification performance relative to the original coefficients. After re-estimation, its combined error increases from 35.4% (at C = 1.81) to 43.1%. This implies that the reduced performance of the Altman Z-score on modern data reflects more than a simple cutoff miscalibration and may instead indicate a genuine decline in the discriminatory relevance of the underlying variable set. The sign reversal of retained earnings ratio X2 is characterised by strong negative skewness (median = -0.10; mean = -4.79). This reflects that sustained accumulated losses have become more common even among healthy firms. The distributional assumptions underlying LDA, particularly multivariate normality, are increasingly violated in the modern Compustat sample (Press & Wilson, 1978). This is further reinforced by the near-collapse of re-estimated coefficient weights across all five Altman variables. When forcing the model to re-fit on modern data, the resulting discriminant boundary between the two groups becomes nearly flat. Thus, the ratio space no longer contains a clean separating base that Altman's original estimates assumed.

The coefficient comparison further reveals a structural shift in determining which financial variables best predict corporate bankruptcy over the common sample period. On the one hand, the leverage variable TLTA, the dominant predictor in Ohlson (1980), declines by 97% in the re-estimated model (from 6.03 to 0.210). On the other hand, INTWO, increases nearly sixfold (from 0.285 to 1.680). A standardised coefficient analysis confirms that these changes reflect genuine structural shifts rather than distributional scaling effects. After adjusting for changes in variable distributions, seven of the nine Ohlson variables still exhibit standardised effect changes exceeding 100%, indicating that the re-estimated model captures a fundamentally different set of bankruptcy predictors. Taken together, these findings propose a structural change in what drives corporate bankruptcy in modern periods. The leverage ratio TLTA has lost discriminatory validity because median leverage among modern non-bankrupt firms (0.52–0.57) now closely approaches the benchmark that characterised bankrupt firms in the 1970s. This leaves the original TLTA variable to no longer be able to effectively separate bankrupt from non-bankrupt firms. Simultaneously, persistent operating losses captured by INTWO, have emerged as the dominant signal. This partially reflects the rise of intangible-asset-intensive, growth-oriented business models whose financial profiles were largely non-existent when the original models were estimated. Fama and French's (2004) paper confirms the structural left-shift in the profitability distribution of newly listed U.S. firms from the 1980s.

This thesis also extends prior research by evaluating model performance across three distinct macroeconomic crisis periods. Contrary to the intuition that systemic financial crises should degrade the performance of accounting-based prediction models, the results show that performance does not deteriorate meaningfully during crisis periods relative to non-crisis times. Under the re-estimated Ohlson model, GFC prediction is marginally easier than during normal times, achieving a combined error of 29.1%, below the normal-period result of 31.2%. This result is consistent with the argument that the firms most severely affected by systemic credit shocks, namely, large, highly leveraged financial institutions, are excluded from the sample during data construction. Moreover, the publicly listed non-financial firms that enter bankruptcy during these periods appear to have experienced financial distress prior to the crisis. The macroeconomic shock only serves as an accelerant rather than an originator of failure.

Nonetheless, the relatively small crisis subsample (N = 9-24) across the crisis analyses limits the statistical power and the findings must be considered with caution.

An industry analysis highlights significant heterogeneity in model performance across the Fama-French ten-industry classification. The re-estimated Ohlson O-score achieves the lowest combined error in seven of ten industries, while the original Altman Z-score dominates in Healthcare (28.4%) and Energy (28.8%), and the Altman Z''-score dominates in Manufacturing (26.0%). The best achievable combined errors range from 25.8% in Business Equipment to 36.4% in Consumer Non-Durables. Ohlson's cross-industry dominance is likely linked to its broader original estimation sample and its logistic regression framework, which is better suited to non-normal, heterogeneous ratio distributions characteristic of diversified modern firm populations. The Altman Z''-score's relative strength in Manufacturing is consistent with its industrial calibration origin, its discriminatory validity is best when firm financial structures most closely resemble those of its original estimation sample. Across all industries, Wholesale and Retail emerges as the most structurally problematic sector, with an average Type I error of approximately 39% across all specifications tested. Traditional ratio-based models are systematically unable to identifying distress in this sector without meaningful adaptation. Consumer Non-Durables and Chemicals exhibit even higher Type I error rates of 57% and 50% respectively under their best-performing specifications. However, given their small holdout bankruptcy counts (N = 7 and N = 2), these figures should be interpreted with caution[8].

This cross-industry heterogeneity is further illustrated by the technology sector, where only one of fifteen predictor variables (INTWO) significantly differentiates bankrupt from non-bankrupt firms, while traditional leverage and profitability ratios carry no discriminatory signal. Thus, tech firms distress appears to follow a structurally distinct trajectory that the existing variable sets are not designed to capture. All in all, these findings emphasise that a single set of model coefficients applied uniformly across sectors will inevitably perform poorly for some groups, highlighting the need for industry-specific prediction models incorporating variables tailored to each industry's financial structure.

Overall, the findings of this thesis contribute to the bankruptcy prediction literature, through a common-holdout design that enables a fair cross-model comparison of Altman and Ohlson, the

---

[8] Their limited sample size prevents any robust conclusion about the structural suitability of traditional models in these sectors.

leverage convergence analysis that documents the structural change in leverage signals of bankruptcy, and the crisis-period and industry-stratified analyses identifying heterogeneity that affect predictive accuracy.

6.1. Practical implications

The practical implications of these findings are clear. The original Ohlson O-score (at its cutoff of 0.038) should not be used in its published form on modern data, as its false-alarm rate of 84.3% makes it operationally impractical. Instead, recalibrating the cutoff alone, reduces the combined error to approximately 34.6% at the optimal threshold (Appendix B2), representing a meaningful improvement at minimal implementation cost. Full re-estimation with updated coefficients achieves a combined error of 30.4%, with the tightest bootstrap IQR of [30.9%, 32.3%] across all specifications. This four percentage point gain from re-estimation over cutoff recalibration alone suggests that practitioners with access to a training sample should re-estimate. For those without, the original Altman Z-score at C = 1.81 represents a fully off-the-shelf alternative at 35.4% combined error with a bootstrap IQR of [32.5%, 34.3%], though it remains meaningfully inferior to the re-estimated Ohlson specification.

In either case, the persistent high Type II error rates across all model variants, ranging from 34% to 57% at their optimal cutoffs, underline that these models are best deployed as first-pass screening and prioritisation tools rather than definitive assessments. They correctly identify approximately 70 to 73% of firms that will file for bankruptcy one year ahead of failure, but should always be complemented by qualitative judgements, market-based signals, and sector-specific expertise. Additionally, practitioners should avoid or be cautious when applying the models for Wholesale or Retail sectors, as their high Type I error, make them not fit for purpose. For corporate governance and audit committees, the finding that INTWO is the most powerful predictor in the re-estimated Ohlson model carries a direct practical implication. The entry into a second consecutive loss year should serve as a trigger for enhanced liquidity monitoring and contingency planning.

6.2. Limitations and further research

This study is subject to several limitations. First, the analysis is restricted to publicly listed U.S. firms in the Compustat sample, which may limit the generalisability of findings to private firms or non-U.S. markets with different bankruptcy regulations and capital structures. Second, the evaluation framework is strictly static, assessing each firm at a single observation point one year before failure. Dynamic hazard models that incorporate the full time-series of financial deterioration (Shumway, 2001) are known to outperform static approaches. Third, the common-sample fairness design, which requires complete data across all three models simultaneously, reduces the serviceable bankrupt universe from 466 to 377 firms, mainly due to the lack of mkvalt (market equity) and CHIN/INTWO. This introduces a selection bias towards more continuously reporting firms, which may potentially overstate model performance relative to real-world conditions. Fourth, the crisis-period subsamples are small with N = 9-24, which limits the statistical power of cross-regime comparisons. Thus, crisis-findings have to be treated with caution. Fifth, the bankruptcy definition used in this study includes both Chapter 7 liquidations and Chapter 11 reorganisations. Previous research suggests that firms entering Chapter 7 generally exhibit significantly more severe pre-bankruptcy financial deterioration than those pursuing Chapter 11 reorganisation (Corbae and D'Erasmo, 2021). Therefore, separating these two failure types as distinct outcomes may yield different model performance or coefficient structures.

Future research could usefully build on these contributions by incorporating sector-specific variables, such as R&D intensity, intangible asset ratios, cash runway metrics, or revenue growth rates, and test whether their inclusion improves discriminatory validity per industry. Additionally, by extending the analysis to international samples would further provide more populous crisis subsets capable of supporting robust cross-regime inference.

Reference

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

Altman, E. I. (2018). A fifty-year retrospective on credit risk models, the Altman Z-score family of models and their applications to financial markets and managerial strategies. *The Journal of Credit Risk*, *14*(4), 1–34. https://doi.org/10.21314/JCR.2018.243

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, *4*(3), 71–111. https://doi.org/10.2307/2490171

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, *1*(4), 267–284. https://doi.org/10.1007/bf00570833

Bozkurt, İ., & Kaya, M. V. (2023). Foremost features affecting financial distress and bankruptcy in the acute stage of COVID-19 crisis. *Applied Economics Letters*, *30*(8), 1112–1123. https://doi.org/10.1080/13504851.2022.2036681

Bris, A., Welch, I., & Zhu, N. (2005). *The costs of bankruptcy* [Working paper]. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=868568

Camska, D., & Klecka, J. (2020). Comparison of prediction models applied in economic recession and expansion. *Journal of Risk and Financial Management*, *13*(3), Article 52. https://doi.org/10.3390/jrfm13030052

Candera, M. (2024). Bankruptcy prediction model in service companies during the COVID-19 pandemic: A comparative analysis. *European Journal of Theoretical and Applied Sciences*, *2*(2), 518–523. https://doi.org/10.59324/ejtas.2024.2(2).45

Collins-Thompson, J., Collins, B., & Maher, J. (2021, November 19). Review of U.S. Business Bankruptcies During the COVID-19 Pandemic. *Federal Reserve Bank of Boston*. https://www.bostonfed.org/publications/supervisory-research-and-analysis-notes/2021/review-of-us-business-bankruptcies-during-the-covid-19-pandemic.aspx

Corbae, D., & D'Erasmo, P. (2021). Reorganization or liquidation: Bankruptcy choice and firm dynamics. *The Review of Economic Studies*, *88*(5), 2239–2274. https://doi.org/10.1093/restud/rdaa091

Corry, C. (2001, January 3). At least 210 net firms failed in 2000. *MarketWatch*. https://www.marketwatch.com/story/at-least-210-net-firms-shut-down-in-2000

Duricova, L., Kovalova, E., Gazdíková, J., & Hamranova, M. (2025). Refining the best-performing V4 financial distress prediction models: Coefficient re-estimation for crisis periods. *Applied Sciences*, *15*(6), Article 2956. https://doi.org/10.3390/app15062956

Dzikowska, M., & Jankowska, B. (2012). The global financial crisis of 2008–2009 and the Fortune Global 500 corporations: Looking for losers among the biggest—Exploratory study. *The Poznań University of Economics Review*, *12*(3), 99–124. https://doi.org/10.18559/ebr.2012.3.852

Fama, E. F., & French, K. R. (2004). New lists: Fundamentals and survival rates. *Journal of Financial Economics*, *73*(2), 229–269. https://doi.org/10.1016/j.jfineco.2003.04.001

Fama, E. F., & French, K. R. (2025). *Kenneth R. French—Data library*. Dartmouth College. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*(2), 179–188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Federal Reserve Bank of St. Louis. (2025). *10-year Treasury constant maturity rate* [Data series DGS10]. FRED. https://fred.stlouisfed.org/series/DGS10

Google. (2026). *Google Scholar* [Database]. https://scholar.google.com/

Grice, J. S., & Ingram, R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research*, *54*(1), 53–61. https://doi.org/10.1016/s0148-2963(00)00126-0

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies*, *9*(1), 5–34. https://doi.org/10.1023/B:RAST.0000013627.90884.b7

Hirakubo, N., & Friedman, H. H. (2002). Dot-bombs: Lessons from the dot-com debacle. *Journal of Internet Commerce*, *1*(2), 89–102. https://doi.org/10.1300/J179v01n02_07

Huo, Y., Chan, L. H., & Miller, D. (2024). Bankruptcy prediction for restaurant firms: A comparative analysis of multiple discriminant analysis and logistic regression. *Journal of Risk and Financial Management*, *17*(9), Article 399. https://doi.org/10.3390/jrfm17090399

Männasoo, K., Maripuu, P., & Hazak, A. (2018). Investments, credit, and corporate financial distress: Evidence from Central and Eastern Europe. *Emerging Markets Finance and Trade*, *54*(3), 677–689. https://doi.org/10.1080/1540496X.2017.1300092

Morris, J. J., & Alam, P. (2008). Analysis of the dot-com bubble of the 1990s. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1152412

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

Press, S. J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, *73*(364), 699–705. https://doi.org/10.1080/01621459.1978.10480080

Roberts, M., & Sufi, A. (2009). Control rights and capital structure: An empirical investigation. *The Journal of Finance*, *64*(4). https://finance.wharton.upenn.edu/~mrrobert/resources/Publications/ControlRightsCapitalStructureJF2009.pdf

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, *74*(1), 101–124. https://www.jstor.org/stable/10.1086/209665

S&P Global Market Intelligence. (2021). *Compustat data guide* [Data guide].

Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, *16*(1), 63–85. https://doi.org/10.1080/07421222.1999.11518234

Tasley, K. (2023, March 14). *US business cycle expansions and contractions*. National Bureau of Economic Research. https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions

U.S. Courts. (2009). *Table F-2: U.S. bankruptcy courts—Business and nonbusiness cases commenced, by chapter of the Bankruptcy Code, during the 12-month period ending December 31, 2009*. https://www.uscourts.gov/sites/default/files/statistics_import_dir/F02Dec09.pdf

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. https://doi.org/10.1007/978-0-387-21706-2

Wharton Research Data Services. (2026). *Compustat North America fundamentals annual* [Dataset]. Standard & Poor's/S&P Global Market Intelligence. https://wrds-www.wharton.upenn.edu/

Appendix

A. Data Construction

A1. Bankrupt per year split by estimation and holdout sample.

Year of Bankruptcy	TOTAL Bankrupt	Estimation B	Hold-Out B	% of Total	% of Estimation	% of Hold-Out
1998	8	8	0	2%	100%	0%
1999	11	6	5	3%	55%	45%
2000	21	12	9	6%	57%	43%
2001	22	15	7	6%	68%	32%
2002	9	8	1	2%	89%	11%
2003	11	4	7	3%	36%	64%
2004	3	1	2	1%	33%	67%
2005	12	7	5	3%	58%	42%
2006	26	17	9	7%	65%	35%
2007	30	21	9	8%	70%	30%
2008	12	6	6	3%	50%	50%
2009	10	8	2	3%	80%	20%
2010	15	7	8	4%	47%	53%
2011	14	7	7	4%	50%	50%
2012	12	7	5	3%	58%	42%
2013	11	8	3	3%	73%	27%
2014	11	8	3	3%	73%	27%
2015	14	7	7	4%	50%	50%
2016	13	9	4	3%	69%	31%
2017	12	8	4	3%	67%	33%
2018	10	7	3	3%	70%	30%
2019	5	3	2	1%	60%	40%
2020	13	7	6	3%	54%	46%
2021	33	20	13	9%	61%	39%
2022	39	18	21	10%	46%	54%
TOTAL	377	229	148			
					45%	

Crisis	Narrow	Wide

A2. Missingness overview

Altman Variable Construction

Variable	# Deleted Obs	% Missing
X1	2.584	2,0%
X2	4.087	3,2%
X3	737	0,6%
X4	17.204	13,3%
X5	628	0,5%
X4_book	143	0,1%

Ohlson Variable Construction

Variable	# Deleted Obs	% Missing
SIZE	-	0,0%
TLTA	-	0,0%
WCTA	2.584	2,0%
CLCA	2.595	2,0%
OENEG	-	0,0%
NITA	631	0,5%
FUTL	1.013	0,8%
INTWO	80	0,1%
CHIN	9.131	7,1%

A3. Analysis of bankrupt firms lost due to common-sample filter.

Bankrupt firms lost to common-sample filter - Analysis

	X1	X2	X3	X4	X4 book	X5	SIZE	TLTA	WCTA	CLCA	OENEG	NITA	FUTL	INTWO	CHIN
Lost bankrupts by missing variable	24	27	39	38	7	39	5	5	24	24	5	39	40	62	76

	Median Tital Asset
Retained bankrupts	128,002
Lost bankrupts	83,544

A4. Correlation matrix of Altman & Ohlson variables (estimation sample)

Altman variables

	X1	X2	X3	X4	X5
X1	1,000	0,715	0,657	0,146	-0,027
X2	0,715	1,000	0,759	-0,032	0,133
X3	0,657	0,759	1,000	-0,027	0,189
X4	0,146	-0,032	-0,027	1,000	-0,237
X5	-0,027	0,133	0,189	-0,237	1,000

Ohlson variables

	SIZE	TLTA	WCTA	CLCA	OENEG	NITA	FUTL	INTWO	CHIN
SIZE	1,000	-0.444	0.433	-0.414	-0.409	0.555	0.34	-0.452	0.012
TLTA	-0.444	1,000	-0.965	0.806	0.652	-0.785	-0.037	0.221	-0.014
WCTA	0.433	-0.965	1,000	-0.844	-0.604	0.776	0.02	-0.194	0.022
CLCA	-0.414	0.806	-0.844	1,000	0.494	-0.662	-0.054	0.208	-0.015
OENEG	-0.409	0.652	-0.604	0.494	1,000	-0.549	-0.066	0.287	-0.016
NITA	0.555	-0.785	0.776	-0.662	-0.549	1,000	0.376	-0.387	0.107
FUTL	0.34	-0.037	0.02	-0.054	-0.066	0.376	1,000	-0.516	0.166
INTWO	-0.452	0.221	-0.194	0.208	0.287	-0.387	-0.516	1,000	-0.046
CHIN	0.012	-0.014	0.022	-0.015	-0.016	0.107	0.166	-0.046	1,000

A5. Auxiliary logistic regression on Altman variables

coefficient	estimate	std_error	z_value	p_value	significance
(Intercept)	0,103	0,178	0,576	0,565	
X1_w	- 0,206	0,203	- 1,014	0,311	
X2_w	0,010	0,008	1,202	0,229	
X3_w	- 0,311	0,193	- 1,610	0,107	
X4_w	- 0,003	0,003	- 0,862	0,389	
X5_w	- 0,096	0,120	- 0,798	0,425	

A6. Variable used in R for firm identification

Overall Company Variable for Identification					
Compustat Variable	**Meaning**	**Periodicity**	**Format**	**Source**	**Document**
gvkey	primary key for each company in the database	Scalar	Number	p. 1038	CDG
cik	identifier for corporations and individuals who	Scalar	Code	p. 467	CDG
conm	Company name	Scalar	Code	p. 512	CDG
conml	Company Legal Name (EDGAR listed)	Scalar	Text	p. 513	CDG
indfmt	Industry Format	Annual	Code	p. 1304	CDG
curcd	ISO Currency Code - Company	Annual	Code	p. 572	CDG
costat	Active/Inactive Status Marker	Scalar	Code	p. 518	CDG
consol	parent and subsidiary accounts are combined or	Annual	Code	p. 514	CDG
popsrc	Population Source	Scalar	Code	p. 3141	CDG
datafmt	Data Format	Annual	Code	p. 591	CDG
datadate	Data Date (Company Annual)	Scalar	Date		CDG
fyear	current fiscal year-end month	Annual	Date	p. 974	CDG
fyr	Fiscal Year-end Month	Annual	Code	p. 975	CDG

A7. Winsorising effect overview per variable

Variable Raw	Variable Winsorised	Values clipped
X1	X1_w	6.456
X2	X2_w	6.202
X3	X3_w	6.482
X4	X4_w	2.442
X4_book	X4_book_w	6.312
X5	X5_w	3.356
SIZE	SIZE_w	6.633
TLTA	TLTA_w	6.532
WCTA	WCTA_w	6.456
CLCA	CLCA_w	6.448
NITA	NITA_w	6.576
FUTL	FUTL_w	6.480
CHIN	CHIN_w	-

B. Results

B1. Raw versus winsorised model performance comparison of all specifications

spec	version	type_I	type_II	combined	sensitivity	specificity
Z_orig_2675	raw	23%	52%	37%	77%	48%
Z_orig_2675	winsorized	23%	52%	37%	77%	48%
Z_orig_181	raw	32%	40%	36%	68%	60%
Z_orig_181	winsorized	31%	40%	35%	69%	60%
Zpp_orig_260	raw	22%	54%	38%	78%	46%
Zpp_orig_260	winsorized	21%	55%	38%	79%	45%
Zpp_orig_110	raw	31%	40%	36%	69%	60%
Zpp_orig_110	winsorized	30%	40%	35%	70%	60%
O_orig_038	raw	3%	84%	44%	97%	16%
O_orig_038	winsorized	3%	84%	43%	97%	16%
O_orig_050	raw	26%	43%	34%	74%	57%
O_orig_050	winsorized	26%	43%	34%	74%	57%

B2. Ohlson O-score cutoff grid: classification performance across probability cutoffs (0-1.0)

Cutoff (Ohlson Model)	Type I (%)	Type II (%)	Combined (%)
0	0.0%	100.0%	50.0%
0.02	2.7%	89.6%	46.1%
0.038	2.7%	84.3%	43.5%
0.05	2.7%	81.7%	42.2%
0.1	8.1%	73.3%	40.7%
0.2	14.9%	62.8%	38.9%
0.4	20.9%	48.3%	34.63%
0.6	**29.7%**	**39.4%**	**34.56%**
0.8	48.0%	31.8%	39.9%
1	100.0%	0.0%	50.0%

B3. Retained earnings ratio (X2) for Amazon Inc. (1997–2007): illustration of structurally negative retained earnings in a financially healthy technology firm (from Compustat)

gvkey	fyear	X2
064768	1998	-0.247
064768	1999	-0.358
064768	2000	-1.08
064768	2001	-1.77
064768	2002	-1.51
064768	2003	-1.36
064768	2004	-0.725
064768	2005	-0.547
064768	2006	-0.421
064768	2007	-0.211

B4. Evolution of median total liabilities to total assets ratio (TLTA) for bankrupt and non-bankrupt non-financial U.S. firms across five-year intervals (1950–2024)

Period	Non-bankrupt firms			Bankrupt firms			Gap (B - NB)
	Median	Mean	N	Median	Mean	N	
Pre-modern era (pre-Ohlson estimation window)							
1950–1954	0,360	0,364	2.226	0,386	0,367	311	0,026
1955–1959	0,353	0,363	2	0,380	0,382	326	0,027
1960–1964	0,401	0,410	7	0,424	0,443	926	0,023
1965–1969	0,466	0,462	12	0,504	0,505	1.598	0,038
Ohlson (1980) estimation window							
1970–1974	0,504	0,504	18	0,556	0,558	2.591	0,052
1975–1979	0,536	0,780	23	0,597	0,705	3.505	0,061
Post-Ohlson era							
1980–1984	0,532	0,613	24	0,618	0,892	3.754	0,086
1985–1989	0,556	1,366	26	0,669	2.203,000	4.239	0,113
1990–1994	0,545	0,842	28	0,673	5.724,000	3.627	0,128
1995–1999	0,522	0,957	35	0,657	5.973,000	3.477	0,135
Modern era (study window)							
2000–2004	0,540	4,087	29	0,686	6.221,000	2.778	0,146
2005–2009	0,525	7,047	24	0,638	2.537,000	2.057	0,113
2010–2014	0,545	5,925	22	0,649	1.584,000	1.468	0,104
2015–2019	0,568	6,420	20	0,684	7.213,000	1.081	0,116
2020–2024	0,549	5,785	19	0,672	2.812,000	713	0,123

B5. Share of technology and intangible firms in Compustat (1950 – 2024)

Includes SIC = 3570–3579 (computers). 3660–3699 (electronics). 7370–7379 (software and data processing)

Period	Narrow tech %	Total firms
1950–1954	3.5%	753
1960–1964	4.2%	2,316
1970–1974	8.2%	5,864
1980-1984	12.5%	7,120
1990-1994	16.5%	8,197
1995–1999	21.8%	9,916
2000–2004	21.8%	8,317
2005-2009	20.1%	6,636
2010–2014	18.7%	6,148
2015–2019	18.0%	5,440
2020-2024	18.3%	4,781

B6. Full bootstrap robustness analysis across 100 resampling iterations for all eight model specifications

Model	Error type	Mean	Median	Q25	Q75	Min	Max	Range Width
Panel A — Altman Z-score (original and re-estimated)								
Z_orig_2675	type_I	20,7%	20,3%	19,4%	22,5%	13,5%	26,4%	3,0%
	type_II	51,9%	51,8%	51,3%	52,5%	49,7%	54,3%	1,2%
	combined	36,3%	36,4%	35,5%	37,2%	32,8%	38,9%	1,7%
Z_orig_181	type_I	27,0%	27,0%	25,7%	28,4%	18,9%	33,1%	2,7%
	type_II	40,1%	40,0%	39,4%	40,6%	37,5%	42,2%	1,2%
	combined	33,6%	33,5%	32,5%	34,3%	29,3%	37,1%	1,8%
Z_re_Cstar	type_I	39,3%	37,5%	31,8%	45,9%	21,6%	68,2%	14,2%
	type_II	43,4%	44,7%	37,0%	50,6%	16,9%	63,9%	13,6%
	combined	41,3%	41,6%	39,3%	43,4%	34,1%	47,3%	4,2%
Panel B — Altman Z''-score (original and re-estimated)								
Zpp_orig_260	type_I	18,2%	17,6%	16,7%	19,6%	12,2%	25,0%	2,9%
	type_II	55,5%	55,5%	54,7%	56,2%	53,1%	57,4%	1,4%
	combined	36,8%	36,7%	36,0%	37,7%	34,1%	39,7%	1,7%
Zpp_orig_110	type_I	24,8%	24,3%	23,6%	26,4%	18,2%	31,8%	2,7%
	type_II	41,8%	41,8%	41,1%	42,4%	39,0%	44,1%	1,3%
	combined	33,3%	33,3%	32,4%	34,2%	29,6%	37,1%	1,8%
Zpp_re_Cstar	type_I	33,4%	33,1%	26,4%	39,2%	14,9%	62,2%	12,8%
	type_II	41,5%	40,9%	37,6%	45,6%	22,8%	59,7%	8,0%
	combined	37,5%	37,1%	35,1%	39,6%	30,7%	44,6%	4,5%
Panel C — Ohlson O-score (original and re-estimated)								
O_orig_038	type_I	**2,9%**	2,7%	2,0%	3,4%	0,0%	5,4%	1,4%
	type_II	84,9%	84,9%	84,4%	85,4%	83,0%	86,5%	1,0%
	combined	43,9%	43,9%	43,5%	44,3%	42,3%	45,3%	0,8%
O_re_Cstar	type_I	23,3%	23,6%	16,2%	28,4%	9,5%	39,2%	12,2%
	type_II	**39,8%**	37,2%	35,1%	46,0%	25,5%	51,0%	10,9%
	combined	31,5%	31,5%	30,9%	32,3%	27,8%	34,8%	1,4%

B7. Standardised coefficient comparison - Ohlson (1980) versus re-estimated model.

Variable	Orig. std. β	Re-est. std. β	Change
TLTA	1.091	0.349	−68%
INTWO	0.058	0.818	+1.310%

OENEG	−0.114	0.239	+310%
NITA	−0.179	0.100	+156%
WCTA	−0.260	0.360	+238%
SIZE	−0.639	0.146	+123%
CLCA	0.056	−0.049	−188%
FUTL	−0.659	−0.007	−99%
CHIN	−0.239	−0.198	−17%

B8. Financial profile of bankrupt firms in crisis versus normal-periods, at t-1, t-2, and t-3 (mean)

	n	X3_w	TLTA_w	NITA_w	INTWO	X1_w
T-1						
Crisis combined	89	-0.247	**0.678**	-0.412	0.719	0.202
Normal	283	-0.446	**1.015**	-0.617	0.749	0.025
T-2						
Crisis combined	78	-0.15	**0.64**	-0.23	0.615	0.225
Normal	245	-0.465	**0.94**	-0.621	0.714	0.044
T-3						
Crisis combined	69	-0.188	**0.646**	-0.322	0.594	0.169
Normal	213	-0.385	**0.844**	-0.518	0.62	0.068

B9. Model performance by crisis period (narrow and wide window)

period	window	spec	type_I_pct	type_II_pct	combined_pct	sensitivity_pct	specificity_pct	n_bankruptcies
Dot-com	narrow	Z_orig_267!	44,4%	57,6%	51,0%	55,6%	42,4%	9
Dot-com	narrow	Z_orig_181	44,4%	44,8%	44,6%	55,6%	55,2%	9
Dot-com	narrow	Z_re_Cstar	44,4%	57,0%	50,7%	55,6%	43,0%	9
Dot-com	narrow	Zpp_orig_2	33,3%	52,3%	42,8%	66,7%	47,7%	9
Dot-com	narrow	Zpp_orig_1	55,6%	38,4%	47,0%	44,4%	61,6%	9
Dot-com	narrow	**Zpp_re_Cst**	33,3%	40,7%	**37,0%**	66,7%	59,3%	9
Dot-com	narrow	O_orig_038	**11,1%**	88,4%	49,7%	**88,9%**	11,6%	9
Dot-com	narrow	O_re_Cstar	11,1%	39,0%	25,0%	88,9%	61,0%	9
Dot-com	wide	Z_orig_267!	31,8%	55,8%	43,8%	68,2%	44,2%	22
Dot-com	wide	Z_orig_181	31,8%	44,3%	38,1%	68,2%	55,7%	22
Dot-com	wide	Z_re_Cstar	31,8%	57,4%	44,6%	68,2%	42,6%	22
Dot-com	wide	Zpp_orig_2	31,8%	56,6%	44,2%	68,2%	43,4%	22
Dot-com	wide	Zpp_orig_1	40,9%	42,1%	41,5%	59,1%	57,9%	22
Dot-com	wide	Zpp_re_Cst	27,3%	42,4%	**34,9%**	72,7%	57,6%	22
Dot-com	wide	O_orig_038	**4,5%**	87,9%	46,2%	95,5%	12,1%	22
Dot-com	wide	O_re_Cstar	13,6%	39,3%	**26,5%**	86,4%	60,7%	22
GFC	narrow	Z_orig_267!	29,2%	51,0%	40,1%	70,8%	49,0%	24
GFC	narrow	Z_orig_181	50,0%	37,9%	43,9%	50,0%	62,1%	24
GFC	narrow	Z_re_Cstar	41,7%	51,8%	46,7%	58,3%	48,2%	24
GFC	narrow	Zpp_orig_2	16,7%	54,0%	35,4%	83,3%	46,0%	24
GFC	narrow	Zpp_orig_1	25,0%	40,4%	32,7%	75,0%	59,6%	24
GFC	narrow	**Zpp_re_Cst**	45,8%	35,1%	**40,5%**	54,2%	64,9%	24
GFC	narrow	O_orig_038	**4,2%**	81,6%	42,9%	95,8%	18,4%	24
GFC	narrow	O_re_Cstar	29,2%	29,0%	29,1%	70,8%	71,0%	24
GFC	wide	Z_orig_2675	32,3%	50,6%	41,4%	67,7%	49,4%	31
GFC	wide	Z_orig_181	48,4%	37,7%	43,0%	51,6%	62,3%	31
GFC	wide	Z_re_Cstar	41,9%	53,3%	47,6%	58,1%	46,7%	31
GFC	wide	Zpp_orig_26	19,4%	53,7%	36,5%	80,6%	46,3%	31
GFC	wide	Zpp_orig_11	29,0%	41,9%	35,5%	71,0%	58,1%	31
GFC	wide	Zpp_re_Csta	41,9%	34,6%	**38,3%**	58,1%	65,4%	31
GFC	wide	O_orig_038	**6,5%**	82,9%	44,7%	93,5%	17,1%	31
GFC	wide	O_re_Cstar	29,0%	32,9%	31,0%	71,0%	67,1%	31

COVID	narrow	Z_orig_267!	0,0%	56,8%	28,4%	100,0%	43,2%	2
COVID	narrow	Z_orig_181	0,0%	43,2%	**21,6%**	100,0%	56,8%	2
COVID	narrow	Z_re_Cstar	0,0%	72,6%	36,3%	100,0%	27,4%	2
COVID	narrow	Zpp_orig_2(	0,0%	65,3%	32,6%	100,0%	34,7%	2
COVID	narrow	Zpp_orig_1:	0,0%	42,1%	21,1%	100,0%	57,9%	2
COVID	narrow	Zpp_re_Cst:	0,0%	43,2%	21,6%	100,0%	56,8%	2
COVID	narrow	O_orig_038	0,0%	88,4%	44,2%	100,0%	11,6%	2
COVID	narrow	O_re_Cstar	0,0%	34,7%	**17,4%**	100,0%	65,3%	2
COVID	wide	Z_orig_267!	9,5%	49,1%	29,3%	90,5%	50,9%	21
COVID	wide	Z_orig_181	28,6%	37,2%	**32,9%**	71,4%	62,8%	21
COVID	wide	Z_re_Cstar	28,6%	66,9%	47,7%	71,4%	33,1%	21
COVID	wide	Zpp_orig_2(	14,3%	55,6%	35,0%	85,7%	44,4%	21
COVID	wide	Zpp_orig_1:	19,0%	39,6%	29,3%	81,0%	60,4%	21
COVID	wide	Zpp_re_Cst:	28,6%	35,2%	31,9%	71,4%	64,8%	21
COVID	wide	O_orig_038	**0,0%**	84,3%	42,2%	100,0%	15,7%	21
COVID	wide	O_re_Cstar	33,3%	36,9%	35,1%	66,7%	63,1%	21
Normal	narrow	Z_orig_267!	20,4%	51,6%	36,0%	79,6%	48,4%	113
Normal	narrow	Z_orig_181	26,5%	39,4%	33,0%	73,5%	60,6%	113
Normal	narrow	Z_re_Cstar	25,7%	57,2%	41,5%	74,3%	42,8%	113
Normal	narrow	Zpp_orig_2(	21,2%	54,4%	37,8%	78,8%	45,6%	113
Normal	narrow	Zpp_orig_1:	30,1%	40,5%	35,3%	69,9%	59,5%	113
Normal	narrow	Zpp_re_Cst:	32,7%	35,3%	**34,0%**	67,3%	64,7%	113
Normal	narrow	O_orig_038	**1,8%**	84,2%	43,0%	98,2%	15,8%	113
Normal	narrow	O_re_Cstar	27,4%	34,9%	31,2%	72,6%	65,1%	113

B10. Model performance by industry period (narrow and wide window)

Ff10 Code	Industry	# Bankrupt Holdout	Model	Type I	Type II	Combined error
1	Consumer Non-Durables	7	Z_orig_2675	42,9%	41,6%	42,2%
1	Consumer Non-Durables	7	Z_orig_181	57,1%	22,7%	39,9%
1	Consumer Non-Durables	7	Z_re_Cstar	57,1%	48,7%	52,9%
1	Consumer Non-Durables	7	Zpp_orig_260	57,1%	42,2%	49,7%
1	Consumer Non-Durables	7	Zpp_orig_110	85,7%	22,7%	54,2%
1	Consumer Non-Durables	**7**	Zpp_re_Cstar	71,4%	20,8%	46,1%
1	Consumer Non-Durables	7	O_orig_038	0,0%	77,9%	39,0%
1	Consumer Non-Durables	7	**O_re_Cstar**	57,1%	15,6%	**36,4%**
2	Consumer Durables	5	Z_orig_2675	20,0%	43,2%	31,6%
2	Consumer Durables	5	Z_orig_181	40,0%	33,0%	36,5%
2	Consumer Durables	5	Z_re_Cstar	40,0%	43,2%	41,6%
2	Consumer Durables	5	Zpp_orig_260	0,0%	48,9%	24,4%
2	Consumer Durables	5	Zpp_orig_110	20,0%	33,0%	26,5%
2	Consumer Durables	5	Zpp_re_Cstar	20,0%	26,1%	23,1%
2	Consumer Durables	5	O_orig_038	0,0%	83,0%	41,5%
2	Consumer Durables	5	O_re_Cstar	0,0%	25,0%	**12,5%**
3	Manufacturing	8	Z_orig_2675	25,0%	47,8%	36,4%
3	Manufacturing	8	Z_orig_181	25,0%	30,1%	27,6%
3	Manufacturing	8	Z_re_Cstar	25,0%	49,1%	37,1%
3	Manufacturing	8	Zpp_orig_260	25,0%	38,4%	31,7%
3	Manufacturing	8	**Zpp_orig_110**	25,0%	27,0%	**26,0%**
3	Manufacturing	**8**	Zpp_re_Cstar	37,5%	21,8%	29,6%
3	Manufacturing	8	O_orig_038	0,0%	83,4%	41,7%
3	Manufacturing	8	O_re_Cstar	37,5%	20,4%	29,0%
4	Energy	9	Z_orig_2675	0,0%	71,8%	35,9%
4	Energy	9	**Z_orig_181**	0,0%	57,7%	**28,8%**
4	Energy	9	Z_re_Cstar	0,0%	87,1%	43,6%
4	Energy	9	Zpp_orig_260	11,1%	66,3%	38,7%
4	Energy	9	Zpp_orig_110	33,3%	45,4%	39,4%
4	Energy	9	Zpp_re_Cstar	11,1%	50,9%	31,0%
4	Energy	9	O_orig_038	0,0%	83,4%	41,7%
4	Energy	9	O_re_Cstar	33,3%	26,4%	29,9%
5	Chemicals	2	Z_orig_2675	50,0%	51,6%	50,8%
5	Chemicals	2	Z_orig_181	50,0%	32,6%	41,3%
5	Chemicals	2	Z_re_Cstar	100,0%	64,2%	82,1%
5	Chemicals	2	Zpp_orig_260	50,0%	44,2%	47,1%
5	Chemicals	2	Zpp_orig_110	100,0%	33,7%	66,8%
5	Chemicals	**2**	Zpp_re_Cstar	100,0%	29,5%	64,7%
5	Chemicals	2	O_orig_038	0,0%	90,5%	45,3%
5	Chemicals	2	**O_re_Cstar**	50,0%	29,5%	**39,7%**
6	Business Equipment	28	Z_orig_2675	21,4%	48,1%	34,8%
6	Business Equipment	28	Z_orig_181	28,6%	37,4%	33,0%
6	Business Equipment	28	Z_re_Cstar	21,4%	50,7%	36,1%
6	Business Equipment	28	Zpp_orig_260	25,0%	49,7%	37,4%
6	Business Equipment	28	Zpp_orig_110	25,0%	41,4%	33,2%
6	Business Equipment	28	Zpp_re_Cstar	39,3%	32,3%	35,8%
6	Business Equipment	28	O_orig_038	7,1%	76,8%	41,9%
6	Business Equipment	28	**O_re_Cstar**	10,7%	40,9%	**25,8%**
7	Telecom	4	Z_orig_2675	0,0%	80,0%	40,0%
7	Telecom	4	Z_orig_181	0,0%	60,0%	30,0%
7	Telecom	4	Z_re_Cstar	0,0%	88,2%	44,1%
7	Telecom	4	Zpp_orig_260	0,0%	75,5%	37,7%
7	Telecom	4	Zpp_orig_110	0,0%	56,4%	28,2%
7	Telecom	**4**	Zpp_re_Cstar	0,0%	52,7%	26,4%
7	Telecom	4	O_orig_038	0,0%	90,9%	45,5%
7	Telecom	4	**O_re_Cstar**	0,0%	49,1%	**24,5%**
8	Wholesale/Retail	24	Z_orig_2675	45,8%	34,3%	40,1%
8	Wholesale/Retail	24	Z_orig_181	50,0%	19,8%	34,9%
8	Wholesale/Retail	24	Z_re_Cstar	54,2%	25,8%	40,0%
8	Wholesale/Retail	24	Zpp_orig_260	29,2%	48,1%	38,6%
8	Wholesale/Retail	24	Zpp_orig_110	41,7%	28,9%	35,3%
8	Wholesale/Retail	24	Zpp_re_Cstar	54,2%	28,0%	41,1%
8	Wholesale/Retail	24	O_orig_038	0,0%	87,4%	43,7%
8	Wholesale/Retail	24	**O_re_Cstar**	45,8%	19,8%	**32,8%**

9	Healthcare	35	**Z_orig_2675**	5,7%	51,1%	**28,4%**
9	Healthcare	35	**Z_orig_181**	11,4%	45,3%	**28,4%**
9	Healthcare	35	Z_re_Cstar	11,4%	65,7%	38,6%
9	Healthcare	35	Zpp_orig_260	5,7%	59,5%	32,6%
9	Healthcare	35	Zpp_orig_110	11,4%	51,1%	31,3%
9	Healthcare	**35**	Zpp_re_Cstar	17,1%	41,2%	29,2%
9	Healthcare	35	O_orig_038	0,0%	86,2%	43,1%
9	Healthcare	35	O_re_Cstar	14,3%	59,0%	36,6%
10	Other	26	Z_orig_2675	30,8%	63,9%	47,3%
10	Other	26	Z_orig_181	50,0%	51,6%	50,8%
10	Other	26	Z_re_Cstar	38,5%	67,5%	53,0%
10	Other	26	Zpp_orig_260	26,9%	67,1%	47,0%
10	Other	26	Zpp_orig_110	38,5%	45,5%	42,0%
10	Other	26	Zpp_re_Cstar	34,6%	46,7%	40,6%
10	Other	26	O_orig_038	7,7%	90,4%	49,1%
10	Other	26	**O_re_Cstar**	34,6%	24,9%	**29,7%**

B11. Bankrupt tech firms vs bankrupt non-tech firms mean variable size

Bankrupt Tech vs. Non-tech firms

variable	pt_mean	nkrupt_mea	welch_t	p_value	significance
X1	0,01	0,07	-0,46	65%	
X2	-13,20	-2,87	-1,22	23%	
X3	-0,52	-0,31	-0,90	37%	
X4	4,16	8,32	-1,08	28%	
X5	1,02	1,10	-0,29	78%	
X4_book	2,50	3,01	-0,30	77%	
SIZE	4,17	4,93	-1,56	13%	
TLTA	1,19	0,86	1,01	32%	
WCTA	0,01	0,07	-0,46	65%	
CLCA	1,32	1,40	-0,18	86%	
OENEG	0,32	0,23	0,99	33%	
NITA	-3,07	-0,47	-1,05	30%	
FUTL	-0,97	-0,55	-0,78	44%	
INTWO	0,89	0,68	2,86	1%	**
CHIN	-0,12	-0,27	1,38	17%	

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/lsm**