Predicting Bankruptcy:
Ask the Employees*

John Knopf and Kristina Lalova‡

Current Version: January 2025

Abstract

The purpose of the paper is to show how employees' attitudes predict bankruptcy throughout various years (phases) of the bankruptcy process – from two and three years before (1), one year before (2), and from the time of filing to the time of liquidation/reorganization (3). We find that our prediction model, inclusive of employees' attitudes, more accurately predicts bankruptcy two to three years before bankruptcy filings, while the other models are more accurate in the year prior to bankruptcy. While already-established statistical models' predictive power is higher (than our model) one year before bankruptcy filings, our model's predictive power is higher (than statistical models) two and three years before bankruptcy filings. Moreover, the addition of employee satisfaction into already-established models improves their predictive performance. We create CatBoost, Random Forest, Support Vector Machine, and Logistic Regression models, and an Autoencoder Anomaly Detection model (GenAI model), all consisting of reviews, and show that textual reviews are in themselves strong predictors of bankruptcy. In survival analyses, we show that employee satisfaction (both in terms of aggregated and individual rating categories) one, two, and three years before bankruptcy filings is a strong predictor whether a company would emerge from bankruptcy successfully. Our paper is the first paper to show that not only is employee satisfaction a predictor of bankruptcy in addition to financial and market data, but that it also is a more powerful predictor of bankruptcy emergence than financial and market data.

Keywords: bankruptcy filing and emergence prediction models, employees' attitudes, in and out-of-sample tests, survival analyses, CatBoost, Random Forest, SVM, Autoencoder Anomaly Detection, Glassdoor
JEL Codes: G33, G41, C53

* We are grateful to Professors Jaideep Shenoy, Yiming Qian, and Lingling Wang (dissertation committee members), Professor Shawn Mobbs, Yumeng Gao, Ding Du, Nur Al Faisal, Austin Starkweather, Hatem Ben Ameur, Bernadette Minton, Andrei Simonov, Stas Sokolinski, Morad Zekhnini, William Organek, Horst Eidenmueller, Vincent Buccola, Daniel Rettl, Su Wang, Alvin Velazquez, and other participants in the FMA Annual Meeting, New Ideas Session (2022), in the World Finance Conference (2023), in the FMA Annual Meeting (2023), Michigan State University Brown Bag (2024), FMA European Conference (2024), MFS Conference (2024), and FMA Asia/Pacific Conference (2024), Harvard-Wharton Insolvency and Restructuring Conference (2024), Conference on Empirical Legal Studies at Emory University School of Law (2024), and Generative AI in Finance Conference (2024). We would also like to thank the Harvard Law School Bankruptcy Roundtable, the Oxford Business Law Blog, and the Reorg Blog for featuring our research – we are very grateful. In addition, we would like to thank Andrew Chamberlain (now at Udemy), Aaron Terrazas (previously at Glassdoor), and Glassdoor for allowing us to use and quote Glassdoor data in this paper. We would also like to thank Dexin Zhou for advising on matching Glassdoor data to other databases. I acknowledge that errors, if any, are mine and all views are mine.
‡ John Knopf, University of Connecticut (john.knopf@uconn.edu), and Kristina Lalova, Michigan State University (lalovakr@msu.edu)

I. Introduction

The purpose of the paper is to test the predictive performance of established bankruptcy prediction models against a new model inclusive of employee information. We track employees' attitudes from before bankruptcy filings to after bankruptcy filings and test their predictability of bankruptcy across the three phases of the bankruptcy process – two and three years before (1), one year before (2), and from the time of filing to the time of liquidation/reorganization (3). We show that before bankruptcy filings, employees have insider information on where the company is headed. Well before bankruptcy or even negative financial performance, managers and employees may be aware of significant problems within their companies. Although managers may be reluctant to disclose this information, employees may reveal problems through dissatisfaction with their jobs and the firm, although we don't necessarily know what the underlying cause for their dissatisfaction is and how they are going to express their dissatisfaction. This employee insider information, however, shows up as a more powerful predictor of bankruptcy, in comparison to financial statement data, two and three years prior to bankruptcy filings. In the year before the bankruptcy, employee insider information is overwhelmed by financial statement data. From the time of bankruptcy filing to the time of liquidation/restructuring, employee insider information overwhelms financial and market data to predict whether the company will emerge from bankruptcy. We empirically test our model for predictability, not causation, similar to what prior bankruptcy prediction literature does. Whether employees are less satisfied because of an impending bankruptcy or whether employee satisfaction impacts the chances of bankruptcy is an interesting topic for further studies.

We document that employee satisfaction is a strong predictor of bankruptcy in all three phases of the bankruptcy process – from two and three years before, one year before, and while in bankruptcy. Specifically, we find that the employee satisfaction model predicts bankruptcy more accurately than any of the existing financial information-based models in all years other than the year immediately prior to a bankruptcy filing. While already-established models' predictive power increases the closer we get to bankruptcy filings, our model's predictive power, in comparison to other models, is higher the further we move from bankruptcy filings. With this finding, we improve on one of the drawbacks of already-established models, more specifically that already-established models don't predict bankruptcy as accurately when moving years back from bankruptcy filings. We additionally find that close to the bankruptcy filing date, models with inclusion of both financial statement and employee satisfaction data outperform models with inclusion of financial data only, according to Adjusted R-Squared, ROC curves, optimal threshold points with the highest sensitivity and specificity, and Type I, Type II, and Total classification error rate analysis. Separately, we hypothesize that if a company is more likely to emerge from bankruptcy, employees are more likely to keep their jobs and be more satisfied with their jobs which is reflected in higher chances of such companies surviving bankruptcy and emerging from it. We document that employee satisfaction predicts bankruptcy emergence and that companies with higher employee satisfaction are more likely to emerge from bankruptcy.

We split our results' analyses into three main sections – testing bankruptcy prediction models with information one year before bankruptcy filings (1), testing Altman's (the most accurate statistical model with financial information one year before) and employees' model with information two and three years before bankruptcy filings (2), and testing a bankruptcy emergence model (3). In the first section of the paper, we test four key bankruptcy models from the literature using a dataset from 2008 to 2020 to show that each one contains unique information regarding the probability of bankruptcy filings. We also build a new model to reflect employees' attitudes before bankruptcy filings and include key variables from each of the four already-established bankruptcy models in the literature in our model. We perform several analyses including parameter estimation tests, bankruptcy classification rate tests, out-of-sample analyses, and boosting machine learning analyses. We compare models based on model-fit criteria, such as the Adjusted R-Squared, Receiver Operating Characteristics area (ROC area), the optimal cutoff point between sensitivity and specificity, and classification error rates.

To go into further detail of the first section of our results, we compare our model with other models one year prior to bankruptcy filings. While Altman's model exhibits the highest ROC and optimal cutoff point between sensitivity and specificity one year prior to bankruptcy filings, when we add the employee satisfaction variable to the other models, the ROC as well as the optimal cutoff point increase. The higher threshold level reflects the fact that the models with the inclusion of employee satisfaction are superior in discriminating between healthy and unhealthy companies. In addition, when we add employee satisfaction to the other four models, their predictive power based on classification error rates improves. Our model still, however, has the highest out-of-sample forecast accuracy, although in in-sample tests, we show that financial and market data outperform employee satisfaction. To go into further detail, we compare our model with Altman's model up to three years prior to bankruptcy filings. Our model exhibits a more consistent ROC over the three years prior to bankruptcy filings and higher ROC two and three years prior to bankruptcy filings in comparison to the Altman's model. Our model exhibits the lower Type I and Type II percentage errors as well. We make the conclusion that employee satisfaction shows up as a predictor of bankruptcy prior to financial statement data, but that it doesn't outperform financial statement data in in-sample tests one year prior to bankruptcy filings. The results signify our expectations that by the time of one year before bankruptcy, companies' prospects and health are already assimilated into their value by market participants and that employee insider information seems to be overwhelmed by financial statement information.

In additional set of tests, we build boosting models using employee satisfaction ratings and employee satisfaction reviews separately and together in the same model to test the predictive performance of employee information alone without any financial statement information added onto the models. For the model with employee satisfaction reviews, we use 13,349 reviews in the year before the bankruptcy filing which include 6,906 actual bankrupt companies' reviews and 6,443 pseudo bankrupt companies' reviews. We find that the model with textual reviews provides an accuracy of 73.37% and a standard deviation of 1.16%. The ROC curve area comes out to be 0.75 which is higher than the ROC curve area for the model using employee satisfaction ratings of 0.64. The accuracy of the model with ratings is also lower and standard deviation is higher (61.66% and 2.37%, respectively). Still, the model with both ratings and reviews exhibits the highest accuracy and the lowest standard deviation. The results point to employee satisfaction textual reviews possessing superior information to employee satisfaction ratings one year before bankruptcy filings.

In the second section of the paper, we more closely compare Altman's model and our model. We find that the employee satisfaction model classifies 76.89%, 76.52%, and 71.53% of bankruptcies one, two, and three years before actual bankruptcy filings at a threshold of 0.5. We also find that the model has sensitivity of 86.54%, 81.65%, 89.91% one, two, and three years before actual bankruptcy filings and specificity of 62.15%, 68.69%, 43.46% one, two, and three years before actual bankruptcy filings indicating that the model has the lower percentage of Type I and Type II errors. We also report sensitivity analyses and provide the accuracy of the employee satisfaction model based on the various rating categories, at a given threshold level, in predicting defaults along with sensitivity and specificity. We are using the 0.5 threshold and report findings for one, two, and three years prior to bankruptcy filings. The results with the breakdown of ratings are consistent with the main findings and we find that the models exhibit a high correct classification rate and lower occurrence of Type I and Type II errors. The results in this section signify our expectation that employees sense issues within their companies two and three years before bankruptcy, although the market doesn't always have knowledge of those issues, and that this employee insider information overwhelms financial statement information.

We also perform additional survival analyses to examine predictions of companies' emergence from bankruptcies and determine that companies with high employee satisfaction are more likely to emerge from a bankruptcy based on mean of sum of various rating categories and individual rating categories. In our sample, we have 301 company reorganizations (Chapter 11 filings) and 26 company failures (Chapter 7 filings). We use a Cox proportional hazards model to test whether employee satisfaction, financial and market data increase the hazard of bankruptcy emergence. We test whether the company would emerge from bankruptcy given their time of filing to their time of liquidation/restructuring. Our results show that for each one-unit increase in employee satisfaction one, two, and three years before bankruptcy filings, the hazard of bankruptcy emergence increases by a factor of 1.57, 1.56, and 1.64, respectively, holding all else constant. Our results show that for each one-unit increase in compensation and benefits and culture and values one year before bankruptcy filings, the hazard of bankruptcy emergence increases by a factor of 1.42 and 1.46, respectively, holding all else constant. Individual rating categories based on information from two years before bankruptcy filings are the most powerful in predicting bankruptcy emergence, although we still document that individual rating categories from one year before and three years before are predictors of bankruptcy emergence.

Overall, our paper provides novel insights into how employee satisfaction shows up as a predictor of bankruptcy in various years (phases) of the bankruptcy process. Since prior literature has proved that companies' financial and market information can predict bankruptcy, we don't contradict other papers' findings, but rather improve on other papers by arguing that employee satisfaction shows up as a predictor of bankruptcy prior to financial and market information, that employee satisfaction improves other models' performance in the year prior to the bankruptcy (where their predictive power is the highest), and that employee satisfaction is a more powerful predictor of bankruptcy emergence than financial and market information. The paper shows that two and three years before bankruptcy employees sense issues within their companies, while one year before bankruptcy financial and market information already reflect what employees knew two and three years prior. From the time the company files for bankruptcy to the time the company either liquidates or restructures, employee satisfaction reflects whether the company would emerge from bankruptcy or not. The results signify that employees hold information on companies' financial health and prospects throughout various years of the bankruptcy process, but that information is expressed in a different way throughout the bankruptcy process.

II. Literature Review and Hypotheses Development

Our paper adds to the literature on bankruptcy prediction models. Finance literature has determined several bankruptcy prediction models, such as Altman's Model (1968), Ohlson's Model (1980), Zmijewski's Model (1984), and Shumway's Model (2001). Since the development of those models, researchers have made efforts to develop models with even greater predictive performance. Prior bankruptcy prediction models have employed financial ratios from financial statements before the bankruptcy filing, while more recent models have used financial market data, such as excess stock returns and stock return volatility, along with the application of the Black-Scholes option-pricing model. The earliest studies on bankruptcy prediction have utilized univariate analyses which have focused on individual ratios and comparison of ratios of failed companies with those of successful firms. Those earliest studies have laid the groundwork for multivariate studies. Among the univariate studies, those to be highlighted include Merwin (1942), Chudson (1945), and Beaver (1966). In his study of small manufacturers, Merwin (1942) found that three ratios are significant indicators of business failure – Net Working Capital to Total Assets, Net Worth to Total Debt, and the Current Ratio – even four or five years before failure. Chudson (1945) has tried to determine whether there is a normal pattern to predict bankruptcy and has reported that there is no such pattern, but there is a clustering of ratios within industry, size, and profitability groups. The most popular univariate study comes from Beaver (1966) in which he compares 30 ratios of 79 failed and 79 non-failed companies in 38 industries and tests their predictive abilities by classifying them under bankrupt and non-bankrupt firms. The author has given future researchers the idea to consider multiple ratios altogether as they might have higher predictive ability than single ratios.

The most popular multivariate study remains Altman (1968). In this paper, the author uses a multivariate discriminant analysis to develop a five-factor model to predict bankruptcy of manufacturing firms. The paper documents that the Z-Score predicts bankruptcy if the firm's score falls within a certain range. The model has high predictive accuracy one year before the merger, but the accuracy falls off two, three, four, and five years before business failure. Since Altman's study on bankruptcy prediction, models' number and complexity have increased. More recent papers have utilized logit and probit models, neural networks, and multivariate discriminant analysis, or some combination of those models. For example, Mensah (1983) has used both multivariate discriminant analysis and logit analysis to predict bankruptcy. In the 1980s, neural networks became the dominant method used. Neural networks analyze inputs to find patterns and develop a model capable of a decision-making process. During the training mode, in which the network learns the decision-making process, several sample cases are run, while during the testing mode the neural network model is validated using hold-out sample data.

The four models tested in the model are Altman (1968), Ohlson (1980), Zmijewski (1984), and Shumway (2001). Altman (1968) uses a multivariate discriminant analysis to develop a five-factor model to predict bankruptcy of manufacturing firms. Altman (2000) adds on his previous paper by showing applications of the Z-score model to private companies, non-manufacturing entities, and to also refer to a new bond-rating equivalent model for emerging markets' corporate bonds and updates the results of the 1968 paper to include results up to 1999. Ohlson (1980) uses a logit model which generates the O-score and shows that four types of variables are predictive of bankruptcy – the company's size, the company's performance, the company's liquidity, and the company's financial structure. Zmijewski (1984) uses a probit model and finds that three variables are predictive of bankruptcy – return on assets, financial leverage, and liquidity. Shumway (2001) uses a hazard model that can capture changes in the company's characteristics over time and shows that market variables are also predictive of bankruptcy. We use key variables from all these various models in addition to an employees' satisfaction proxy in a hazard model and compare the predictive performance of already-established models in their original settings with the predictive performance of our model. We compare the predictive performance of all models both in in-sample and out-of-sample tests. In addition, we compare the predictive performance of the best-performing models – the Altman's model and the employees' model – one, two, and three years before bankruptcy filings. Our approach is described in the data and methodology sections.

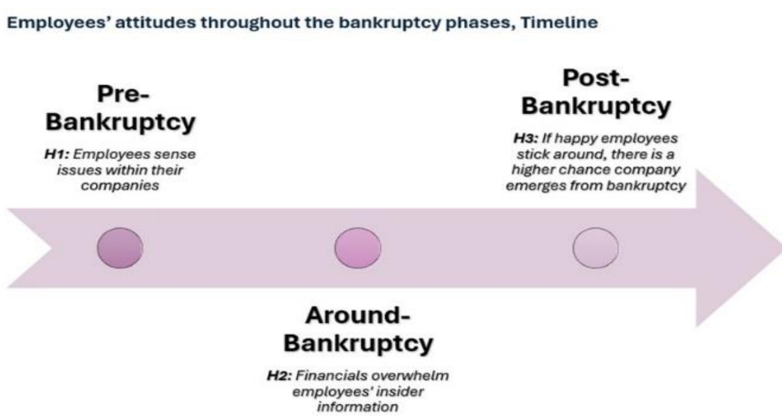
More recent papers on bankruptcy prediction have reviewed the literature on bankruptcy prediction and have compared various models' predictive power utilizing only financial and market variables. Charitou et al. (2004) use an artificial neural network for bankruptcy prediction and credit analysis. As mentioned above, neural networks are designed to find patterns to solve given problems. The authors provide evidence that machine learning methods can provide results as good as established statistical methods. Altman et al. (2014) provide a review of the bankruptcy prediction literature and test the Z-Score model in an international setting to show that the inclusion of country-specific variables could improve the predictive performance of the Z-Score model. Jones et al. (2017) provide a wide range of classifiers for analysis to predict corporate bankruptcies and show that when using financial information, the traditional classifiers perform well, but modern classifiers are recommended because of their performance in cross-sectional and longitudinal tests. In the spirit of those papers, we complement our statistical analysis with the application of a machine learning method with employee textual reviews to show that reviews are also predictive of bankruptcy years before bankruptcy filings and that reviews provide additional predictive power to ratings.

Other recent papers have used other modern machine learning techniques. For instance, Barboza et al. (2017) use data from 1985 to 2013 on U.S. companies analyzing more than 10,000 firm-year observations and compare the predictive performance of financial ratios using machine learning techniques, such as random forest, bagging, and boosting. Their research contributes to the debate regarding the superiority of computational methods over established statistical methods. Our paper is also closely related to papers that use more modern techniques of bankruptcy prediction. For instance, Hillegeist et al. (2004) test market variables in Black-Scholes-Merton option-pricing model; Alfaro et al. (2008) test accounting variables in neural networks and boosted decision trees; Chen et al. (2010) test accounting and market variables in a KVM model; Olson et al. (2012) test accounting and market variables in decision trees, neural networks, and support vector machines; Tobback et al. (2017) test accounting and relational variables in a smoothed wvRN and support vector machines; Zelenkov and Volodorskiy (2021) test accounting and macroeconomic variables in machine learning models.

Another strand of literature is Glassdoor literature, which focuses on various themes. For example, papers have examined companies' corporate climate and its impact on SEC fraud enforcements, its impact on the return effect, its impact on merger and acquisition synergies, and its impact on private equity deals (Ji et al. (2017); Green et al. (2019); Lalova (2023); Lalova (2023)). To be more specific, Lalova (2023) shows that M&A deals with high similarity between acquirer and target employee morale achieve greater short-run and long-run synergies and that high target employee morale is contagious for acquirer employees, while Lalova (2023) shows that low similarity between private equity acquirers and their targets enhances post-merger integration and long-run synergies. Marchetti (2019) measures organizational culture using text reviews employees have posted on Glassdoor and finds that pre-deal acquirer-target culture compatibility is positively associated with the likelihood of a deal announcement and with superior stock returns. Chen et al. (2022) propose an employee sentiment index, complementing investor sentiment and manager sentiment indices, and find that high employee sentiment predicts low monthly (weekly) market returns significantly both in- and out-of-sample. Campbell and Shang (2021) show that information extracted from Glassdoor reviews can be used to develop measures with useful properties for measuring misconduct risk and those measures clearly discriminate between high and low misconduct firms. Liu et al. (2022) use Glassdoor data to show that firms offer higher quality maternity leave benefits in labor markets where female talent is relatively scarce. Welch and Yoon (2021) find that high-ability managers allocate resources to ESG efforts in a way that enhances shareholder value. In the spirit of those papers, I use Glassdoor data to show that employees can predict bankruptcy filings and emergences.

While there is no direct evidence suggesting that employee satisfaction can predict company failures, employee satisfaction is a factor that can impact employees' productivity, creativity, and innovation, which in turn could impact companies' financial health. In this paper, we define employee satisfaction as employees' attitudes toward and perceptions of the tasks employees have in the companies they work for and various firm dynamics. Those perceptions toward various firm dynamics include perceptions toward career opportunities, compensation and benefits, culture and values, work-life balance, senior leadership, and overall organizational performance. All these firm dynamics are determinants of employee satisfaction. To put it in other words, leadership styles (senior leadership), organizational culture (culture and values), work-life balance (work-life balance), compensation and benefits (compensation and benefits), and opportunities for career advancement (career opportunities) are all determinants of employee satisfaction. We take the level of perceptions toward those various firm dynamics to build employee satisfaction. While employee satisfaction can influence bankruptcy risk through its effect on organizational performance, it is not the sole predictor of bankruptcy. As mentioned, researchers have determined that financial and market information is a predictor of bankruptcy in various models and settings. The downside of those various papers is that the financial and market models' predictive power decreases the further we move from bankruptcy filings. Thus, in this paper, we don't contradict other papers' findings, but rather improve on other papers by arguing that employee satisfaction shows up as a predictor of bankruptcy prior to financial and market information, that employee satisfaction improves other models' performance in the year prior to bankruptcy (where their predictive power is the highest), and that employee satisfaction is a more powerful predictor of bankruptcy emergence than financial and market information.

We expect to find that employees sense significant issues within their companies, which materialize into negative and poor attitudes toward their employers two and three years before bankruptcy filings. Given the market might not always see the full picture of companies' financial health, we expect to document that this employee insider information overwhelms financial statement information two and three years before bankruptcy filings. However, as companies move closer to bankruptcy filings, the true value of companies is revealed to the market, which we expect to be expressed in financial statement information overwhelming employee insider information. Finally, as companies have filed for bankruptcy, we expect employees' attitudes to reveal their employees' belief whether financials show us the full picture of whether the company would emerge (survive) bankruptcy. If happy and satisfied employees have stuck around from three years before bankruptcy filings to the time of the actual filing, we expect the company to have a higher chance of surviving bankruptcy. Below is a timeline of employees' attitudes throughout the bankruptcy process.



In this paper, we show that employees have information on upcoming bankruptcies two and three years before bankruptcy filings by testing various already-established models in the literature against a newly-established one consisting of employees' attitudes and show that its predictive power is greater than that of the established models up to two and three years before bankruptcy filings, and that when a proxy for employee satisfaction is added to the established models, their predictive power increases. With these results, we contribute to the bankruptcy prediction literature whose downside is not being able to show the same bankruptcy predictive power two and three years prior to bankruptcy filings as one year prior to bankruptcy filings. Our paper is also the first paper that shows how employees' information predicts bankruptcy across various phases of the bankruptcy prediction process. Building on prior bankruptcy prediction and Glassdoor research, we build the three following hypotheses:

i. Employee satisfaction shows up as a predictor of bankruptcy prior to financial statement data.

ii. Around bankruptcy announcements, financial information overwhelms employee satisfaction information, but the addition of employee satisfaction into already-established models improves their predictive performance.

iii. Employee satisfaction around the bankruptcy filing predicts whether the company will emerge from bankruptcy.

III. Data and Methodology

In this paper, we utilize UCLA's LoPucki, Glassdoor, Compustat, and CRSP databases. We use UCLA's LoPucki data for bankruptcy filings' information, Glassdoor data for employees' satisfaction metrics, Compustat data for financial statement variables, and CRSP data for return variables. Our sample consists of 327 bankruptcy filings in the period between 2008 and 2020. In addition to including data on employees' satisfaction, we include profitability variables, liquidity variables, leverage variables, firm-size variables, other firm characteristics, such as standard deviation of the return, excess return, and logarithm of firm age, in accordance with prior bankruptcy prediction studies. The variables we choose to include in our model are based on key variables included already in prior bankruptcy prediction literature. An alternative to determine the variable to be included would be to utilize Principal Component Analysis to select variables with the highest explanatory power. For the construction of the control sample, we use a matched-pair methodology to compare the estimation procedures of the models we are testing. We construct a matched-pair sample based on industry, size (based on a company's total assets), and book-to-market. The match generated for each bankrupt firm is based on minimizing the absolute value of the ratio of the difference between the firm size and book-to-market of the bankrupt firm and that of the healthy firm to the firm size and book-to-market of the bankrupt firm. The control group is matched with a 1:1 ratio via the nearest neighbor based on the minimization technique described above. A failed company's total assets and book-to-market one year prior to the bankruptcy filing year are used to match with the total assets and book-to-market of a healthy company. The approach of matching in a year when the two companies are healthy allows us not to overfit our data (Lennox, 1999). The models we include in our study are Altman (1968), Ohlson (1980), Zmijewski (1984), Shumway (2001), and we build a new model where we use some key variables from previous models and include the employees' satisfaction proxy.

- Altman (1968): Multiple discriminant analysis with financial ratios

$Z = \beta' X$

where Z is the MDA score, and X represents the variables listed. Cutoff point: Z ≥ 2.675, classified as non-bankrupt. Under X, we include Net Working Capital/Total Assets, Retained Earnings/Total Assets, EBIT/Total Assets, Market Value of Equity/Book Value of Total Liabilities, and Sales/Total Assets. Altman's model is a linear discriminant analysis model which uses the above-mentioned ratios to predict the likelihood of a company facing financial distress. The model assigns weights to several financial ratios based on their predictive power and combines them to generate a single score, which is known as the Z-Score. The score is then used to classify companies into different bankruptcy risk categories. The model is still widely used, despite some of its limitations, as long as the assumptions of linearity and equal covariances are reasonably met.

- Ohlson (1980): Logit model with financial ratios

$P = (1 + exp\{-\beta' X\})^{-1}$

where P is the probability of bankruptcy and X represents the variables listed. The logit function maps the value of $\beta' X$ to a probability bounded between 0 and 1. Under X, we include Ohlson's Size, Total Liabilities/Total Assets, Working Capital/Total Assets, Current Liabilities/Current Assets, OENEG (a dummy variable indicating if Total Liabilities exceed Total Assets), Operating Income/Total Liabilities, INTWO (a dummy variable indicating if Net Income was negative for the last 2 years, and 0 otherwise), and Change in Net Income. The Ohlson's O-Score model uses both financial ratios and accounting variables to calculate a bankruptcy prediction score for companies. The model assigns weights to each variable based on their statistical significance in predicting bankruptcy. Based on their score, companies are classified into two categories – those in high risk of distress and those in low risk of distress.

- Zmijewski (1984): Probit model with financial ratios

$P = \Phi(\beta' X)$

where P is the probability of bankruptcy and X represents the variables listed, and $\Phi(\bullet)$ represents the cumulative normal distribution function. The probit function maps the value of $\beta' X$ to a probability bounded between 0 and 1. Under X, we include Net Income/Total Liabilities, Total Liabilities/Total

Assets, and Current Assets/Current Liabilities. The model is one of the early models of bankruptcy prediction based on financial ratios. The model's coefficients are estimated using maximum likelihood estimation based on financial information for both distressed and non-distressed companies. The model makes an estimate of how much each financial ratio contributes to the risk of bankruptcy while controlling for the effects of other financial ratios in the model. Once the coefficients have been estimated, the model is used to predict the probability of bankruptcy for new firms. Companies are then classified into two categories of distressed and non-distressed companies. Some of the benefits of the model are its simplicity and interpretability, while some of its drawbacks include limited predictive power, lack of dynamic factors, and linear assumptions.

- Shumway (2001): Hazard model with both financial ratios and market data

$$P_{i,t} = (1 + \exp\{-y_{i,t}\})^{-1}$$

$$y_{i,t} = \alpha + \beta' X_{i,t-1} = \beta' \begin{bmatrix} X_{1,t-1} \ldots X_{1,t-j} \\ \cdots \cdots \\ X_{n,t-1} \ldots X_{n,t-j} \end{bmatrix}$$

where *P* is the probability of bankruptcy and *X* represents the variables listed. This is a multi-period logit model, but instead of treating each firm year as an independent observation, all prior values of the independent variables for a particular firm are included in the information set. *n* represents the number of independent variables, and *j* represents the number of time periods prior to time *t* for which data are available. Under X, we include Net Income/Total Liabilities, Total Liabilities/Total Assets, Relative Size, Cumulative Annual Return in prior year minus the value weighted CRSP index in prior year, and Standard Deviation of Return in prior year. The technique behind Shumway's model involves estimating bankruptcy probability over a specific period based on financial ratios. Some of the benefits of Shumway's model include its accuracy and its flexibility allowing for incorporation of time-varying predictors and censoring of data (which enables the model to capture any changes in the companies' financial health over time). We refine the model slightly by including one firm-year observation for each healthy company and all firm-year observations for failing companies.

- Employees' Model: Hazard model with employees' data

$$P_{i,t} = (1 + \exp\{-y_{i,t}\})^{-1}$$

$$y_{i,t} = \alpha + \beta' X_{i,t-1} = \beta' \begin{bmatrix} X_{1,t-1} \dots X_{1,t-j} \\ \dots \dots \\ X_{n,t-1} \dots X_{n,t-j} \end{bmatrix}$$

where *P* is the probability of bankruptcy and *X* represents the variables listed. This is a multi-period logit model. *n* represents the number of independent variables, and *j* represents the number of time periods prior to time *t* for which data are available. We keep only one firm-year observation for each bankrupt firm but all firm-year observations for surviving and pseudo firms. We go back three years before bankruptcy filings and track all variables included in our model up to three years of bankruptcy filing. Under X, we include Mean of Employees' Satisfaction, EBIT/Total Assets, Working Capital/Total Assets, Total Liabilities/Total Assets, Change in Net Income, Excess Return, Standard Deviation of Return, and Logarithm of the closing price of prior year. We include liquidity, profitability, and leverage variables, as well as a firm size variable, in addition to market variables, such as the company's lagged return and volatility. We are expecting that companies with higher liquidity and profitability are less likely to go bankrupt, while we are expecting that companies with higher leverage are more likely to go bankrupt. We follow Shumway's approach of a multi-period hazard model where we include one firm-year observation for bankrupt companies and all firm-year observations for non-bankrupt companies. Because of the inclusion of many serially correlated variables in the same model, we need to adjust standard errors where the Wald-Chi Square needs to be divided by the number of firm-years.

IV. Results

[Insert Table 1 here]

Table 1 presents summary statistics for each variable used in the study over the period 2008-2020. The dataset contains 327 actual bankrupt firms, which we define as non-healthy, and 327 non-bankrupt firms, which we define as healthy. Panel A of Table 1 groups the variables into various categories going back three years before actual bankruptcy filings. The profitability variables include EBITTA (EBIT to Total Assets), Sales, NITA (Net Income to Total Assets), and CHIN (Change in Net Income), and they measure the ability of the firm to generate sufficient profits or returns to remain a company that is going concern. The liquidity variables include WCTA (Working Capital to Total Assets), CLCA (Current Liabilities to Current Assets), FUTL (Funds from Operations to Total Liabilities), INTWO (a dummy variable signifying if the firm has a negative Net Income over the past two years), and they measure the ability of the firm to meet its short-term obligations. The leverage variables include RETA (Retained Earnings to

Total Assets), MVETL (Market Value of Equity to Total Liabilities), TLTA (Total Liabilities to Total Assets), OENEG (a dummy variable indicating if Total Liabilities exceed Total Assets), and they measure the relative amount of debt and other obligations of the firm. The firm-size variables include OhlsonSize (calculated as the natural logarithm of Total Assets to the GNP Price-Level index) and RelativeSize (calculated as the natural logarithm of the Number of Outstanding Shares multiplied by the Year-End Share Price to Total Market Value), and they reflect the ability of larger firms to trade through difficult times and their ability to be less likely to go bankrupt. We also include other firm characteristics, such as the Excess Return, Standard Deviation of Excess Return, and we determine that bankrupt firms tend to have lower and more volatile returns and that they have substantially underperformed in their years prior to bankruptcy. In addition, those firms tend to be younger (as observed by their firm age). Panel B of Table 1 reports the mean and standard deviation of the sum of all rating categories up to three years before actual bankruptcy filings. As can be seen, the mean increases the closer we get in time to the bankruptcy filing and the standard deviation increases the closer we get in time to the bankruptcy filing. Panel C of Table 1 presents the breakdown of the sum of mean and standard deviation into different categories – *Career Opportunities*, *Compensation Benefits*, *Culture Values*, *Senior Leadership*, *Work-Life Balance*, and *Overall Rating* – up to three years before actual bankruptcy filings. One can observe that the rating category means are relatively low on a possible range between 1 and 5. Figure A1 presents the correlation heat map of correlations between the Z-Score and the mean and standard deviation of sum of rating categories, and the individual employees' satisfaction rating categories. The correlation between the Z-Score and the mean of the sum and the various rating categories is low and negative one and two years before actual bankruptcy filings, while the correlation between the Z-Score and the mean of the sum and the various rating categories is low and positive three years before actual bankruptcy filings. This could be attributed to the fact that the predictive power of the Z-Score decreases the further we move back in time from actual bankruptcy filings. This finding is documented by Altman (1968), but it is also what we find in our sample of bankruptcies between 2008 and 2020. Our main findings are discussed in the following few pages.

[Insert Figure 1 and Figure 2 here]

Figure 1 shows the yearly distribution of our bankruptcy sample from 2008-2020. As expected, 2009 and 2020 are the years with the most bankruptcies (63 and 51, respectively). Figure 2 shows a plot of the average marginal effects and the predictive margins with a 95% confidence interval of the mean of sum of rating categories one, two, and three years before bankruptcy filings. Average marginal effects show how changes in the level of employee satisfaction affect the probability of bankruptcy. The marginal effect of level of employee satisfaction indicates that the increase in the level of employee satisfaction decreases the probability of bankruptcy. Predictive margins show how the probability of bankruptcy varies across different levels of employee satisfaction. The higher the level of employee satisfaction, the lower the average predicted probability of bankruptcy. Both figures show that those companies with lower employee satisfaction are more exposed to the risk of an actual bankruptcy filing.

IV.I. Results with Information from One Year Before Bankruptcy

[Insert Table 2 here]

In this Results section, we test established statistical models against our model with information from one year before bankruptcy filings. Table 2 reports the parameter estimates from the various four models and our model discussed above. We use the same variables that have been used by the original authors in the authors' proposed setting except for Altman's model in which we use a multi-period logistic setting. We include one firm-year observation for each company filing for bankruptcy and all firm-year-observations for matched control and surviving companies for all three years prior to bankruptcy filing. The discriminant analysis coefficients (both standardized and unstandardized) are presented in Table A2 columns (1) and (2) in the Appendix. The estimated coefficients in Table 2 from the models are bolded if there is any statistical significance. We set the dependent variable to be equal to 1 for bankrupt firms and to 0 otherwise. Therefore, if we observe a positive coefficient, it means that a higher value of that variable increases the likelihood of bankruptcy. We can see that companies with lower profitability and liquidity, but higher leverage are more likely to file for bankruptcy, and smaller firms are more likely to file for bankruptcy. Our employee satisfaction model includes employee satisfaction proxy and profitability, liquidity, and leverage variables additionally. It also includes the lagged return and volatility of the return, and the firm-size variable as indicated by the company's stock price. If we compare the models based on Adjusted R-Squared, which is a more accurate measure of the models' fit than the Pseudo R-Squared, Altman's model has the highest Adjusted R-Squared (0.3710), Ratings' model has the second-best Adjusted R-Squared (0.2780), and Ohlson's model has the third best Adjusted R-Squared (0.1790). We further examine whether the models' fit improves with the addition of employee satisfaction to the models in Table 3.

[Insert Figure 3 and Figure 4 here]

In Figure 3, one can observe that Altman's model has the highest ROC (0.9084), while the employees' satisfaction model is the second-best performing model one year before bankruptcy filings (0.8867). We further examine how the ROC changes from three years to one year before bankruptcy filings for the Altman's and employees' satisfaction models in the following paragraphs. Figure 4 shows the optimal cutoff point for the various models and the relationship between sensitivity and specificity. Sensitivity is the correct classification of true default and specificity is referred to as the correct classification of true non-default. The two types of uncertainty are referred to as Type I and Type II errors. The optimal cutoff point is where the sensitivity and specificity are the highest. The analysis shows that the Altman's model has the highest optimal threshold level, but that our model outperforms all other models based on the optimal threshold level characteristic other than the Altman's model. The optimal threshold for both models is between 0.60 and 0.75, which is where we have the highest sensitivity and specificity. The higher threshold level for these two models reflects that these models are superior in discriminating between healthy and unhealthy firms. Overall, the results in the above tables show that Altman's and our model have the highest predictive performance, although Altman's model slightly outperforms our model. We move on to show how the addition of employee satisfaction proxy into the other models impacts their predictive performance. We then take the two best-performing models – Altman's model and our model – to compare their predictive performance one, two, and three years before bankruptcy filings.

[Insert Table 3 here]

[Insert Figure 5 and Figure 6 here]

Table 3 presents parameter estimates for the four models with inclusion of the employee satisfaction proxy, such as the approach in Table 2. We use the same variables that have been used by the original authors in the authors' proposed setting except for Altman's model in which we use a multi-period logistic setting. The discriminant analysis coefficients (both standardized and unstandardized) are presented in Table A2, columns (3) and (4), in the Appendix. With addition of the employee satisfaction proxy, Adjusted R-Squared for Altman's model increases to 0.4160 (from 0.3710) and Adjusted R-Squared for Ohlson's model increases to 0.2330 (from 0.1790). We can conclude that addition of employee satisfaction to the models improves the models' fit and their predictive performance (which adds proof to our second hypothesis). While Table 2 proves that Altman's model has the highest predictive performance (which proves that right before the bankruptcy filing, financial information overwhelms any other information about failing companies), Table 3 proves that still the addition of the employee satisfaction proxy adds predictive power to the other models. In addition, we examine whether addition of employee satisfaction to the models improves their receiver operating characteristics. Figure 5 shows a plot of the ROC curves for the four models with inclusion of the employee satisfaction proxy. The ROC with the employee satisfaction proxy increases in comparison to the ROC of the models without the employee satisfaction proxy. The Altman's ROC increases from 0.9084 to 0.9182. The greatest improvement in the ROC, however, is in Zmijewski's model – the ROC increases from 0.5391 to 0.5959. Figure 6 plots the optimal cutoff point between sensitivity and specificity for the models in Table 3. The threshold and optimal cutoff point for the Ohlson's model increase with the addition of the employee satisfaction variable into the model, while the threshold and optimal cutoff point for the Altman's model decrease with the addition of the employee satisfaction variable into the model. The higher threshold level reflects that Ohlson's model with employee satisfaction is superior in discriminating between healthy and unhealthy firms. This points to employee satisfaction improving the predictive performance of the Ohlson's model based on the optimal threshold characteristic.

[Insert Table 4 here]

[Insert Figure 7 here]

We move onto reporting results of bankruptcy classification rates and out-of-sample forecast accuracy. In Tables 4, 5, 6, and 7, we use all actual model econometric techniques and variables each author proposed in each respective paper. In Table 4, we report the rate of Types I and II errors for the five models tested in the paper. The table shows the incidence of Type I errors (classifying a bankrupt firm as healthy) and Type II errors (classifying a healthy firm as bankrupt), according to model scores. For example, under the Altman model, the classification total error rate is minimized at the 50th percentile where the Type I error rate is 20.18% and the Type II error rate is 16.75% with a Total error rate of 36.93%. Under the employee satisfaction model, the classification total error rate is minimized at the 50th percentile where the Type I error rate is 37.31% and the Type II error rate is 14.49% with a Total error rate of 51.79%. Although the employee satisfaction model exhibits a lower Type II error rate, the Altman's model still exbibits a lower Type I error rate and a lower Total error rate. We can confirm that the Altman's model is more powerful in predicting bankruptcy based on classification error rates. Still, the employee satisfaction model is the second most powerful model to predict bankruptcy based on classification error rates. Figure 7 plots the Total error rate based on different percentiles. Both the classification rate plot for Altman's model and the classification rate plot for the employee satisfaction model follow the same behavior which is also reflected in the results in Table 4.

[Insert Table 5 here]

[Insert Figure 8 here]

In Table 5, we report the rate of Types I and II errors for the four models with the inclusion of the employee satisfaction proxy. The addition of employee satisfaction improves the classification error rates of the Altman's model and the Ohlson's model. For instance, at the 50th percentile, where classification error rates are minimized without the employee satisfaction proxy, the Altman's model exhibits a Type I error rate of 17.43% and a Type II error rate of 13.71% with a Total error rate of 31.14%, which represents a 2.75% decline in the Type I error rate, a 3.05% decline in the Type II error rate, and a 5.79% decline in the Total error rate. For instance, at the 50th percentile, where classification error rates are minimized without the employee satisfaction proxy, the Ohlson's model exhibits a Type I error rate of 29.97% and a Type II error rate of 19.16% with a Total error rate of 49.13%, which represents a 27.73% decline in the Type I error rate, a 6.08% increase in the Type II error rate, but still an overall decline of 21.75% in the Total error rate. Figure 8 shows the classification rates for the four models with inclusion of employee satisfaction. Altman's and Ohlson's classification rate curves behave consistently with and without the employee satisfaction proxy.

[Insert Table 6 here]

The out-of-sample performance of each model is summarized in Table 6. We estimate the bankruptcy probability for each observation in the out-of-sample period. These observations in the out-of-sample period are ranked by the estimated bankruptcy probability, and we group observations into deciles on this basis. The actual number and predicted number of bankruptcies along with the Chi-square statistics from among those firms are classified in each decile. We use the rolling windows method to do so. Our first estimation is focused on firm-year observations from 2008 to 2009 and bankruptcies in 2010, while the second estimation is focused on firm-year observations from 2008 to 2010 and bankruptcies in 2011. The estimated coefficients used to predict bankruptcies in 2020 are based on firm-year observations from 2008 to 2018 and bankruptcies from 2009 to 2019. The windows continue to expand. We report both the numbers of actual and predicted bankruptcy filings classified into deciles by their estimated probability of bankruptcy. As one can see, the employee satisfaction model predicts bankruptcy more accurately than Altman's model. The Chi-Square statistics across the various deciles of the employee satisfaction model are lower and more consistent than that of the Altman's model. The predictions for the employee satisfaction model most closely match the actual bankruptcies. We conclude that the employee satisfaction model exhibits the highest out-of-sample forecast accuracy one year before bankruptcy filings, although Altman's model still exhibits the highest in-sample performance. We further conduct an out-of-sample performance analysis for the various models with the inclusion of the employee satisfaction proxy in those models.

[Insert Table 7 here]

Table 7 reports the out-of-sample performance for the various models with the inclusion of the employee satisfaction proxy. Similar to what we do in Table 6, in Table 7, we report the actual number of bankruptcies from among the firms classified in each decile with our estimations being based on the rolling-windows method. Like in Table 6, we report the actual and predicted numbers for each model. The addition of the employee satisfaction proxy improves the out-of-sample performance in some, but not across all, deciles for Ohlson's, Zmijewski's, and Shumway's models. The actual and predicted bankruptcy filings are identical (51 actual and 52.99 predicted, 51 actual and 51.34 predicted, 52 actual and 52.63 predicted for Ohlson's, Zmijewski's, and Shumway's models, respectively).

IV.II. Results with Information from Two and Three Years Before Bankruptcy

[Insert Table 8 and Table 9 here]

In this Results section, we compare Altman's and employees' models with information from two and three years before bankruptcy. Table 8 reports sensitivity analyses and provides the accuracy of Altman's and employee satisfaction's models, at a given threshold level, in predicting in-sample defaults along with sensitivity and specificity. We use a multi-period logit setting for both models in this table. We include one firm-year observation for each company filing for bankruptcy and all firm-year observations for matched control and surviving companies for all three years prior to bankruptcy filing. In Table A3, we present error rates but in a discriminant analysis setting for both models, and in this setting, we also find more balanced and lower stratified and unstratified total error rates of the ratings' model persisting three years before the bankruptcy filing. We define sensitivity as the true classification of default firms when the firms defaulted, and specificity as the true classification of non-defaulted firms as non-defaulted. The two types of uncertainty are Type I and Type II errors. Type I error classifies a bankrupt company as non-bankrupt, while Type II error classifies a non-bankrupt firm as bankrupt. We can see that Altman's model classifies 70.54%, 61.00%, 63.40% of the out-of-sample bankruptcies one, two, and three years before actual bankruptcy filings at the threshold level of 0.5. Additionally, we can see that the employee satisfaction model classifies 76.89%, 76.52%, and 71.53% of the in-sample bankruptcies one, two, and three years before actual bankruptcy filings at the level of 0.5. We can also see that the model has sensitivity of 86.54%, 81.65%, 89.91% one, two, and three years before actual bankruptcy filings and specificity of 62.15%, 68.69%, 43.46% one, two, and three years before actual bankruptcy filings which shows that the model has a lower sensitivity, but a higher specificity than Altman's model.

In Table 9, we report the sensitivity analyses and provide the accuracy of the employee satisfaction's model based on the various rating categories, at a given threshold level, in predicting in-sample defaults along with sensitivity and specificity. We are using the 0.5 threshold and report findings for one, two, and three years prior to bankruptcy filings. The results with the breakdown of ratings are consistent with the main findings in Table 8 and we find that the models with the breakdown of ratings exhibit a high correct classification rate and low occurrence of Type I and Type II errors.

[Insert Figure 9 here]

Figure 9 plots the ROC curve from the tests in Table 8. The ROC score for Altman's model is 0.9084, 0.7808, 0.7075 one, two, and three years before bankruptcy filings, while the ROC score for the employee satisfaction model is 0.8867, 0.8280, 0.8212 one, two, and three years before bankruptcy filings. We can conclude that the ROC score for the employees' satisfaction model is more consistent, even though it still decreases the further we go back from the time of the actual bankruptcy filing

IV.III. Results from Machine Learning and Survival Analyses

[Insert Table 10 and Figure 10 here]

In this Results section, we provide machine learning and survival analyses. Table 10 and Figure 10 provide results from CatBoost, Random Forest, Support Vector Machine (SVM), and Logistic Regression model, and an Autoencoder Anomaly Detection model (GenAI model) for predicting bankruptcy by leveraging employee reviews to uncover patterns that signal financial distress in companies, while Table 11 and Table 12 provide results from survival analyses providing insight on whether employee attitudes impact bankruptcy emergence.

The CatBoost model is a gradient boosting algorithm designed to handle categorical features and sparse data, making it particularly well-suited for textual data like employee reviews. It builds an ensemble of decision trees sequentially, where each tree corrects the errors of its predecessors by minimizing a loss function. CatBoost is robust against overfitting due to its advanced regularization techniques and incorporates a unique algorithm to process categorical data efficiently, which is crucial for features, such as employee reviews. In this context, CatBoost leverages the TF-IDF-transformed text data to capture complex, non-linear relationships, allowing it to identify subtle patterns in employee feedback that are predictive of bankruptcy.

The Random Forest model, on the other hand, is an ensemble learning algorithm that constructs multiple decision trees during training and aggregates their predictions to improve robustness and accuracy. Each tree is trained on a random subset of the data, with the final prediction determined by majority voting, ensuring reduced overfitting compared to single decision trees. Random Forest is highly effective with high-dimensional data and provides interpretability through feature importance scores. While it may not handle non-linear relationships as intricately as CatBoost, it excels in producing stable and reliable results, making it a strong benchmark for bankruptcy prediction.

The Support Vector Machine (SVM) model is a powerful algorithm designed to find the optimal hyperplane that separates data into different classes. For regression or classification, SVM uses kernels like the Radial Basis Function to map data into higher-dimensional spaces, enabling the model to capture complex, non-linear relationships. The model is particularly effective at handling high-dimensional data, as it identifies the decision boundary with the maximum margin between classes. With the addition of probability estimates, SVM further enhances its predictive capabilities by providing insights into the confidence of its predictions. Unlike other ensemble methods, SVM focuses on a subset of the training data (support vectors) to determine the hyperplane, making it robust to overfitting in sparse datasets like textual data. Its ability to effectively separate overlapping classes while minimizing misclassifications makes it highly suitable for nuanced prediction, such as bankruptcy.

The logistic regression model is designed to predict continuous outcomes by identifying relationships between input features and the target variable. Using textual data from employee reviews, the regression model attempts to capture patterns in employee reviews that correlate with bankruptcy risk. By minimizing error metrics, such as Mean Squared Error, during training, the model learns to provide a continuous score reflecting the likelihood or severity of bankruptcy. While effective at handling high-dimensional data, the regression model assumes a linear or near-linear relationship between features and bankruptcy risk, which may limit its ability to fully capture complex, non-linear patterns in the data. This approach provides valuable insights into the relative impact of employee reviews on bankruptcy risk, offering a probabilistic framework for understanding company financial health.

The Autoencoder Anomaly Detection model, which is a type of Generative AI model, is designed to identify anomalies, such as bankruptcy, by reconstructing input data and measuring the reconstruction error. Using textual data from employee reviews, the autoencoder learns to encode and decode the input into a compressed latent representation. During training, the model minimizes reconstruction error metrics, such as Mean Squared Error, on non-bankrupt companies, learning to represent patterns in normal employee feedback effectively. When applied to bankrupt companies, the model often fails to reconstruct their data accurately, resulting in higher reconstruction errors, which serve as a signal for potential anomalies. While effective at capturing high-dimensional and complex relationships, this approach relies on the assumption that bankrupt companies exhibit fundamentally different patterns in their employee reviews compared to non-bankrupt companies. By setting a reconstruction error threshold, the autoencoder provides a probabilistic framework for detecting anomalies, offering valuable insights into bankruptcy risk based on deviations in employee feedback patterns.

We take the following steps to clean and process employee reviews for all models. First, any missing values of employee reviews are filled with an empty string. Thereafter, the text is cleaned by converting the text to lowercase, removing special characters and numbers, removing common stop words, such as “the”, “and”, and “is”, and applying lemmatization to reduce words to their root form. The pros, cons, and feedback sections are merged into the same column for analysis. The processed text is transformed into numerical features using TF-IDF which turns text into numerical values based upon its importance in the text. The number of features is reduced to 5,000 to reduce dimensionality and computational cost. We use the Synthetic Minority Oversampling Technique to oversample the minority class. Since the dataset contains a large number of reviews for non-bankrupt companies, we apply under-sampling to the majority class and limit the number of non-bankrupt company reviews is limited to 22,009 reviews. We then use the SMOTE technique to the training data to create synthetic samples for the minority class which ensures that the model doesn't become biased toward the majority class. This procedure was first introduced by Chawla et al. (2002) and has been used in machine learning research thereafter. It allows us to use random under-sampling to trim the number of examples in the majority class and to oversample the minority class to balance the class distribution. We use the same approach for all models, except for the Autoencoder Anomaly Detection model in which case we train the model on non-bankrupt reviews only. Employee pros, cons, and feedback sections on Glassdoor represent the parameters in our models and we utilize 6,334 failed company reviews (pros, cons, and feedback) and 22,009 non-failed company reviews (pros, cons, and feedback).

The CatBoost model demonstrated exceptional predictive performance in identifying bankruptcy risk using employee review data. The ROC curve (AUC = 0.9249) indicates that the model performs well in distinguishing between bankrupt and non-bankrupt firms. This high AUC suggests that the CatBoost model is adept at capturing patterns in employee textual feedback, pros, and cons. These results of low false positive and false negative rate translate to high accuracy and robust classification performance, with the CatBoost model effectively utilizing the textual features to predict bankruptcy. The Random Forest model also showed strong performance but slightly lagged behind the CatBoost model in terms of predictive power. The ROC curve (AUC = 0.9317) suggests slightly better separation between classes compared to CatBoost, indicating that Random Forest is highly competitive. While the Random Forest model achieves excellent classification, the slightly higher false negatives (1,328) compared to CatBoost suggest it may struggle more with borderline cases.

The ROC curve of the support vector machine model shows an impressive AUC of 0.9502, indicating the model's excellent ability to differentiate between bankrupt and non-bankrupt firms. This demonstrates that the SVM effectively captures patterns in the employee review data, enabling high predictive accuracy. These results indicate that the SVM model excels in minimizing false positives (155) while maintaining a strong recall for bankrupt firms. However, the false negatives (1,393) suggest that some bankrupt companies are being misclassified, which could be a point for further refinement in feature engineering or hyperparameter tuning. In the regression model, if we look at the actual vs. predicted plot, most predicted values are clustered near 0 and 1, aligning closely with the actual values. However, there is a slight spread around the line, suggesting the model has some prediction errors. Outliers can be observed where the predicted values deviate significantly from the actual ones, which indicates areas where the model struggles to capture the underlying patterns. In the residual plot, ideally, residuals should be randomly scattered around the horizontal red dashed line at 0, indicating no systematic errors. Here, the residuals exhibit a clear structure, suggesting the model might not fully capture some patterns in the data.

The Autoencoder Anomaly Detection model demonstrates promising results for identifying bankruptcy risk. The training and validation loss curves indicate effective learning, with both losses decreasing steadily and stabilizing after 10-15 epochs, showing that the model generalizes well to unseen data and avoids overfitting. The reconstruction error distribution further highlights the model's capability to distinguish between bankrupt and non-bankrupt companies. Non-bankrupt companies exhibit lower reconstruction errors, while bankrupt companies generally have higher errors, reflecting their deviation from the patterns learned during training. The red dashed line represents the threshold for classification, separating anomalies (bankrupt companies) from normal data. While the two distributions show a clear separation, some overlap exists, potentially leading to false positives (non-bankrupt companies classified as bankrupt) or false negatives (bankrupt companies classified as non-bankrupt). Overall, the Autoencoder effectively identifies bankruptcy as an anomaly, and the threshold can be adjusted to balance precision and recall based on the business context and tolerance for classification errors.

[Insert Table 11 and Table 12 here]

We also perform additional survival analyses in Tables 11 and 12 to examine predictions of companies' emergence from bankruptcies and determine that companies with high employee satisfaction are more likely to emerge from bankruptcy based on mean of sum of various rating categories and individual rating categories. In our sample, we have 301 company reorganizations (Chapter 11 filings) and 26 company failures (Chapter 7 filings). We use a Cox proportional hazards model to test whether employee satisfaction, financial and market data increase the hazard of bankruptcy emergence. We test whether the company would emerge from bankruptcy given their time of filing to their time of liquidation/restructuring. In the Cox proportional hazards model, the coefficient estimates represent the log hazard ratio which is indicative of the change in the hazard of the event (which in this case is bankruptcy emergence) associated with a one-unit change in the predictor variable, holding all else constant. We document our findings in Tables 11 and 12. For instance, we document positive and statistically significant coefficients for levels of employee satisfaction at the 5% level (as expressed in Mean_1, Mean_2, and Mean_3). If we want to interpret the meaning of those coefficients, we can compute the hazard ratio. We can do that by exponentiating the coefficient. The coefficient for Mean_1 of 0.451 represents a hazard ratio of 1.57; the coefficient for Mean_2 of 0.444 represents a hazard ratio of 1.56; the coefficient for Mean_3 of 0.494 represents a hazard ratio of 1.64. The results show that for each one-unit increase in employee satisfaction one, two, and three years before bankruptcy filings, the hazard of bankruptcy emergence increases by a factor of 1.57, 1.56, and 1.64, respectively, holding all else constant. The mean of rating categories has high predictive power for bankruptcy emergences in all three years before bankruptcy filings, while the mean of various individual categories has the highest predictive power two years before bankruptcy filings. Still, one year before bankruptcy filings, perceptions of culture and compensation matter most for bankruptcy emergences. The coefficients for both compensation and benefits (compensation_mean_1) and culture and values (culture_mean_1) are both statistically significant at the 10% level. If we interpret those coefficients, the coefficient for compensation_mean_1 of 0.352 represents a hazard ratio of 1.42, while the coefficient for culture_mean_1 of 0.378 represents a hazard ratio of 1.46. The results show that for each one-unit increase in compensation and benefits and culture and values one year before bankruptcy filings, the hazard of bankruptcy emergence increases by a factor of 1.42 and 1.46, respectively, holding all else constant. Although the results in this section of the paper are only preliminary, we can conclude that employee satisfaction is predictive of bankruptcy emergence.

V. Conclusion

We show that employee satisfaction is a powerful predictor of bankruptcy two and three years before bankruptcy filings, one year before bankruptcy filings, and from the time of bankruptcy filings to the time of companies' restructuring/liquidation. We document that employees sense difficulties within their companies two and three years before bankruptcy filings, while financial and market information overwhelms employees' insider information one year before bankruptcy filings (likely because of companies' prospects and health already having been revealed to market participants). Once the company files for bankruptcy, employees' insider information up to three years before bankruptcy filings predicts whether the company would emerge from bankruptcy or not. We test four key bankruptcy models from the literature using a dataset from 2008 to 2020 to show that each one contains unique information regarding the probability of bankruptcy filings. We also build a new model to reflect employees' attitudes before bankruptcy filings and include key variables from each of the four already-established bankruptcy models in the literature in our model. The model generated shows that employee satisfaction information shows as a predictor of bankruptcies before financial statement information. Right around bankruptcy filings, financial information overwhelms other information about companies including employee satisfaction information. Still, we show that the addition of employee satisfaction to already-established models improves their predictive performance. We document that employee satisfaction reviews provide information relevant for bankruptcy predictions on top of the information provided by employee satisfaction ratings. We also document that employee satisfaction predicts bankruptcy survival (emergence) using additional analyses to our main analyses. Our findings suggest that employees sense financial difficulties and problems in the companies they work for years before they show up on financial statements. The findings could be beneficial for our understanding of information that can reveal a company's failure in addition to already-established information that bankruptcy prediction models are using. The paper also shows that we need reliable and trustworthy techniques that will improve credit risk analysis.

References

[1]**Alfaro, Esteban, Noelia Garcia, Matias Gomez, & David Elizondo,** 2008. Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems*, 45: 110-122.

[2]**Altman, Edward I.,** 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23: 589–609.

[3]**Altman, Edward I.,** 2000. Predicting financial distress of companies: Revisiting the Z-score and Zeta models. Available at *NYU Stern*: PredFnclDistr.pdf (nyu.edu).

[4]**Altman I. Edward, Malgorzata Iwanicz-Drozdowska, Erkki K. Laitinen, & Arto Suvas,** 2014. Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. Available at *SSRN*: Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model by Edward I. Altman, Malgorzata Iwanicz-Drozdowska, Erkki K. Laitinen, Arto Suvas :: SSRN.

[5]**Barboza, Flavio, Herbert Kimura, & Edward Altman,** 2017. Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83: 405-417.

[6]**Beaver, William H.,** 1966. Financial ratios as predictors of failure. *Journal of Accounting Research*, 71–111.

[7]**Campbell, Dennis & Ruidi Shang,** 2022. Tone at the bottom: Measuring corporate misconduct risk from the text of employee reviews. *Management Science*, 68(9): 7034-7053

[8]**Charitou, A., E. Neophytou, & C. Charalambous,** 2004. Predicting corporate failure: Empirical evidence for the UK. *European Accounting Review*, 13(3): 465-497.

[9]**Chawla, N. V., K. W. Bowyer, L. O. Hall, & W. P. Kegelmeyer,** 2002. SMOTE: Synthetic minority over-sampling technique.

[10] Chen, Xiaohong, Xiaoding Wang, & Desheng Dash Wu, 2010. Credit risk measurement and early warning of SMEs: An empirical study of listed SMEs in China/ *Decision Support Systems*, 49(3): 301-310.

[11]**Chen, J., Guohao Tang, Jiaquan Yao, & Guofu Zhou,** 2023. Employee sentiment and stock returns. *Journal of Economic Dynamics & Control*, 149.

[12]**Chudson, Walter A.,** 1945. A survey of corporate financial structure. Volume Title: The pattern of corporate financial structure: a cross-section view of manufacturing, mining, trade, and construction. *NBER*.

[13]**Green, T. Clifton, Ruoyan Huang, Quan Wen, & Dexin Zhou,** 2019. Crowdsourced employer reviews and stock returns. *Journal of Financial Economics*, 134(1): 236-251.

[14]**Hillegeist, Stephen A., Elizabeth K. Keating, Donald P. Cram, & Kyle G. Lundestedt,** 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies*, 9: 5-34.

[15]**Ji, Yuan, Oded Rozenbaum, & Kyle Welch,** 2022. Corporate culture and financial reporting

risk: Looking through the Glassdoor. Available at *SSRN*: Corporate Culture and Financial Reporting Risk: Looking Through the Glassdoor by Yuan Ji, Oded Rozenbaum, Kyle Welch :: SSRN.

[16]Jones, Stewart, David Johnstone, & Roy Wilson, 2017. Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting*, 44(1-2): 3-34.

[17]Lalova, Kristina, 2023. The value of employee morale in mergers and acquisitions: Evidence from Glassdoor. Available at *SSRN*: The Value of Employee Morale in Mergers and Acquisitions: Evidence from Glassdoor by Kristina Lalova :: SSRN.

[18]Lalova, Kristina, 2023. Untangling employee morale in private equity deals. Working paper.

[19]Lennox, C., 1999. Identifying failing companies: A re-evaluation of the logit, probit, and DA approaches. *Journal of Economics and Business*, 51(4): 347-364.

[20]Liu, Tim, Christos Makridis, Paige Ouimet, & Elena Simintzi, 2023. The distribution of nonwage benefits: Maternity benefits and gender diversity. *The Review of Financial Studies*, 36(1): 194-234.

[21]Marchetti, Arianna, 2019. Firms of a feather merge together: The coordination benefits of compatible cultures. Available at *SSRN*: Firms of a Feather Merge Together: The Coordination Benefits of Compatible Cultures by Arianna Marchetti :: SSRN.

[22]Mensah, Yaw M., 1983. The differential bankruptcy predictive ability of specific price level adjustments: some empirical evidence. *The Accounting Review*. Vol. LVIII, No. 2.

[23]Merwin, 1942. Financing small corporations in five manufacturing industries, 1926-36. *NBER*.

[24]Ohlson, James, 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18: 109–131.

[25]Olson, David L., Dursun Delen, & Yanyan Meng, 2012. Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2): 464-473.

[26]Shumway, Tyler, 2001. Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business*, 74: 101–124.

[27]Tobback, Ellen, Tony Belotti, Julie Moeyersoms, Marija Stankova, & David Martens, 2017. Bankruptcy prediction for SMEs using relational data. *Decision Support Systems*, 102: 69-81.

[28]Welch, Kyle & Aaron Yoon, 2021. Do high-ability managers choose ESG projects that create shareholder value? Evidence from employee opinions. *Review of Accounting Studies*.

[29]Zelenkov, Yuri & Nikita Volodarskiy, 2021. Bankruptcy prediction on the base of the unbalanced data using multi-objective selection of classifiers. *Expert Systems with Applications*, 185.

[30]Zmijewski, Mark E., 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Suppl.), 59–82

Table 1: Summary Statistics

Table 1 presents summary statistics for the full bankruptcy sample from 2008 through 2020. The final dataset contains 327 bankruptcies and 327 non-bankrupt firms. EBITTA = earnings before interest and taxes to total assets; Sales = sales to total assets; NITA = net income divided by total assets; CHIN = change in net income from year of to year before; WCTA = working capital to total assets; CLCA = current liabilities to current assets; FUTL = income from operations after depreciation divided by total liabilities; INTWO = 1 if net income was negative for the previous two years, 0 otherwise; RETA = retained earning to total assets; METL = market equity to total liabilities; TLTA = total liabilities to total assets; OENEG = 1 if total liabilities exceed total assets, 0 otherwise; TLMTA = total liabilities to market value of total assets; Ohlson's Size = log(total assets/GNP price-level index), the index assumes a base value of 100 for 1968; Relative Size = log(the number of outstanding shares multiplied by year-end share price then divided by total market value); Price = log of closing price at end of previous fiscal year; SDReturn = standard deviation of excess return; LagExReturn = lagged excess return; FirmAge = log(years for which firm has traded). Panel A presents the above-mentioned bankruptcy models' characteristics one, two, and three years before. Panel B presents mean and variability summary statistics for the 327 bankruptcies, while Panel C presents mean and variability of breakdown of various rating categories one, two, and three years before bankruptcy filings.

1.1. Bankruptcy Models' Characteristics One, Two, and Three Years Before

Variable	Observations	Mean	Standard Deviation
Profitability Variables			
EBITTA_1	654	-0.07530	0.325769
Sales_1	528	1936.804	8580.106
NITA_1	654	-0.14671	0.344535
CHIN_1	654	-0.25264	7.626269
Liquidity Variables			
WCTA_1	654	0.003879	0.398063
CLCA_1	654	1.129634	2.739787
FUTL_1	654	-0.10756	0.549372
INTWO_1	654	0.401109	0.490577
Leverage Variables			
RETA_1	654	-0.609070	1.894407
MVETL_1	654	0.791590	2.604564
TLTA_1	654	0.628457	0.554343
OENEG_1	654	0.164510	0.371081
Firm-Size Variables			
OhlsonSize_1	654	1.601003	1.857073
RelativeSize_1	654	-1.58703	1.561818
Other Firm Characteristics			
Price_1	654	0.736131	1.237024
ExcessReturn_1	654	-0.22312	0.644702
SDReturn_1	654	0.036120	0.035597
FirmAge_1	654	1.361375	1.276400

1.2. Mean and Variability Statistics

Variable	Observations	Mean	Standard Deviation
Mean_1	327	7.825330	9.150692
SD_1	327	2.210938	3.388707
Mean_2	327	6.310011	8.681792
SD_2	327	1.974335	3.281781
Mean_3	327	4.521087	7.666320
SD_3	327	1.351704	2.782840

2.3. Rating Categories' Statistics

Variable	Observations	Mean	Standard Deviation
One Year Before			
careeropps_mean_1	327	2.698533	0.873781
careeropps_sd_1	327	0.806543	0.634286
compensation_mean_1	327	3.204054	1.019952
compensation_sd_1	327	0.693903	0.563472
culture_mean_1	327	2.355393	1.496977
culture_sd_1	327	0.742689	0.737253
seniorleadership_mean_1	327	2.615828	1.022545
seniorleadership_sd_1	327	0.896511	0.693012
worklife_mean_1	327	3.160145	1.022187
worklife_sd_1	327	0.855969	0.714880
overall_mean_1	327	3.025266	0.875284
overall_sd_1	327	0.824228	0.599971
Two Years Before			
careeropps_mean_2	262	2.783246	0.870941
careeropps_sd_2	262	0.893290	0.599946
compensation_mean_2	262	3.304863	0.746865
compensation_sd_2	262	0.821785	0.566678
culture_mean_2	262	2.291046	1.435886
culture_sd_2	262	0.882462	0.717721
seniorleadership_mean_2	262	2.609627	0.981455
seniorleadership_sd_2	262	0.962891	0.653779
worklife_mean_2	262	3.114457	0.864257
worklife_sd_2	262	0.889338	0.623642
overall_mean_2	262	3.091540	0.798091
overall_sd_2	262	0.930298	0.589306
Three Years Before			
careeropps_mean_3	271	2.590691	0.847830
careeropps_sd_3	271	0.827568	0.589574
compensation_mean_3	271	3.068088	0.978264
compensation_sd_3	271	0.714481	0.550266
culture_mean_3	271	2.186082	1.338209
culture_sd_3	271	0.722396	0.663047
seniorleadership_mean_3	271	2.441988	0.876578
seniorleadership_sd_3	271	0.862535	0.638214
worklife_mean_3	271	3.008461	0.866640
worklife_sd_3	271	0.844089	0.606252
overall_mean_3	271	2.950795	0.875191
overall_sd_3	271	0.886155	0.582013

Figure 1: Number of Bankruptcies per Year

Figure 1 presents the number of bankruptcies per year from UCLA's LoPucki data over the paper sample from 2008 through 2020.

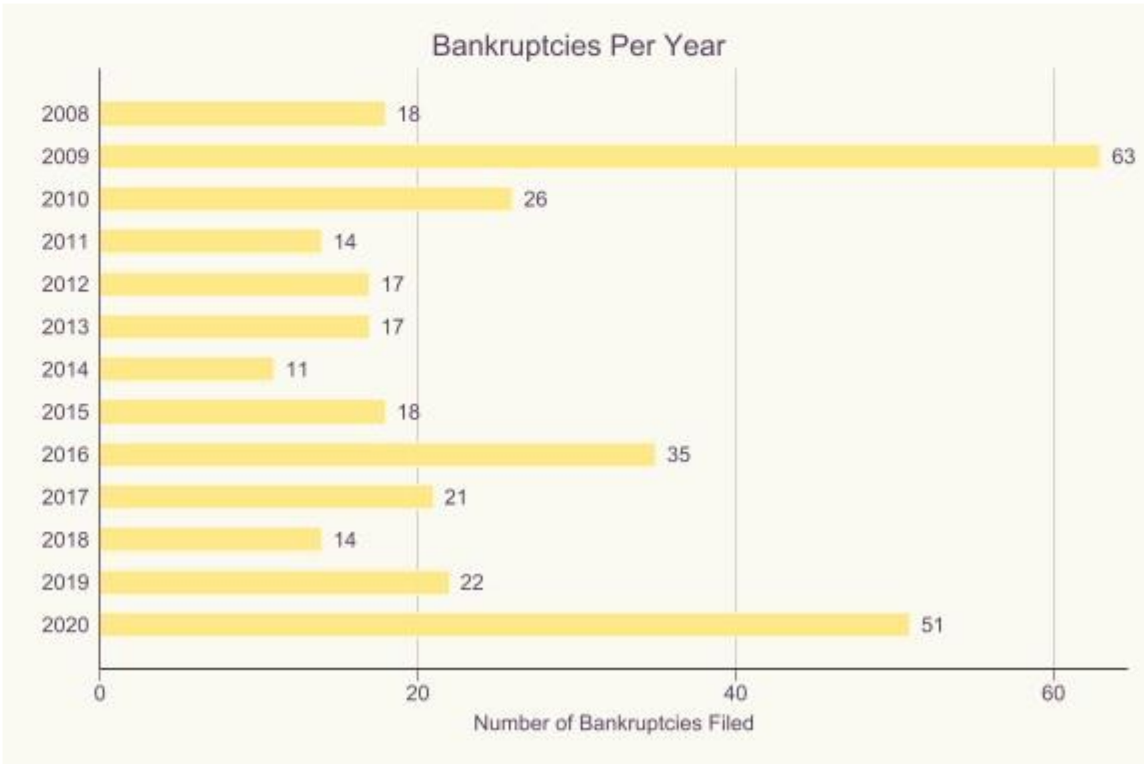
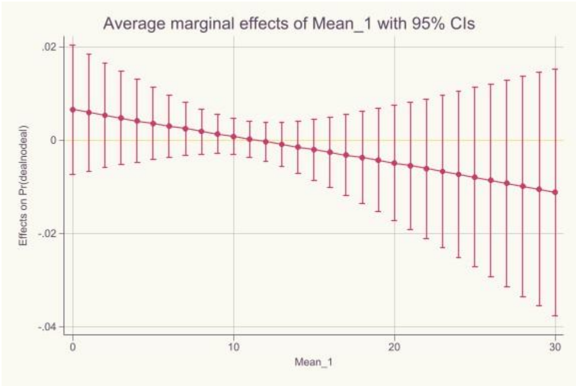
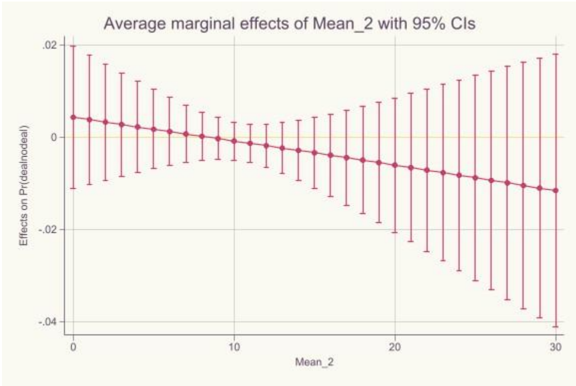


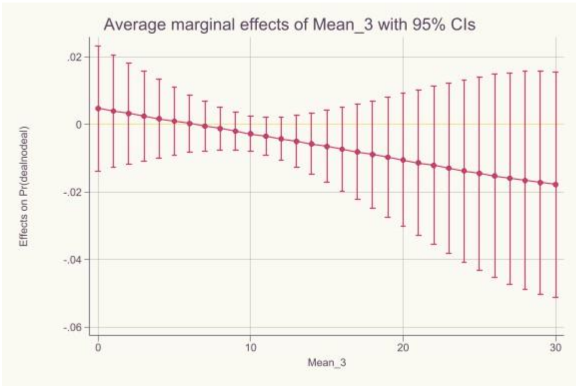
Figure 2: Marginal Effects and Predictive Margins of Mean Ratings

Figure 2 shows a plot of the average marginal effects and the predictive margins with 95% confidence interval of the mean of sum of rating categories one, two, and three years before bankruptcy filings. The figures plot the predicted outcome at various levels of the mean one, two, and three years before.

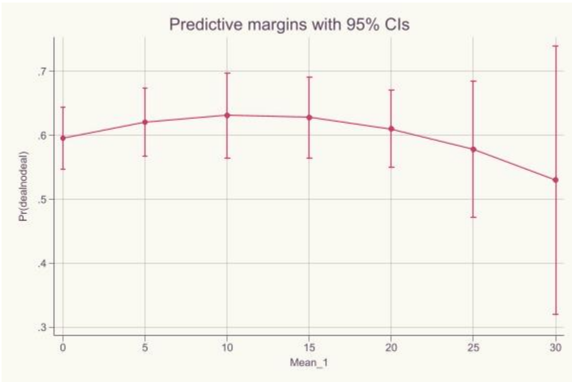
2.1. Average Marginal Effects

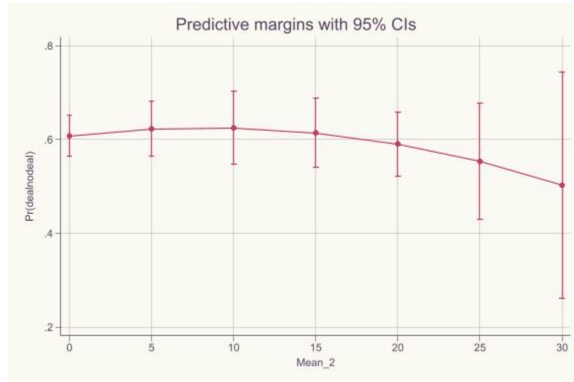






2.2. Predictive Margins





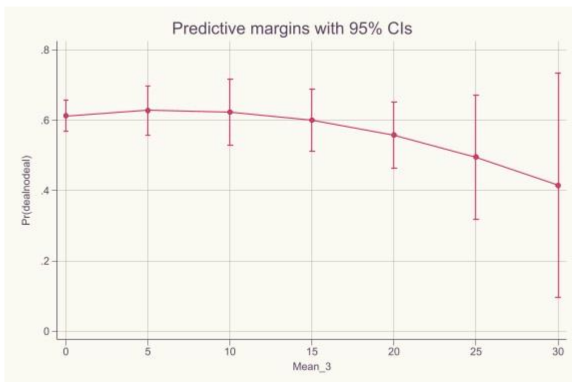


Table 2: Estimation Results for Bankruptcy Models

Table 2 presents the parameter estimates from various bankruptcy prediction models. The sample is from 2008 to 2020 and contains 327 bankruptcies and 327 non-bankrupt firms. Bold font signifies an estimate that is statistically significant. The adjusted Wald Chi-Square statistics from the logistic regression are presented adjacent to each parameter estimate. The Pseudo R-Squared for each model is reported. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Ohlson		**Altman**		**Zmijewski**	
	Estimate	Wald-Chi Square	Estimate	Wald-Chi Square	Estimate	Wald-Chi Square
Intercept	**0.827*** **	58.46	**1.309*** **	31.29	0.195	17.66
Mean_1						
EBITTA_1			-1.452 (1.804)	0.648		
Sales_1			0.0000299 (5.68e-05)	0.278		
NITA_1	-0.775 (1.215)	0.407			-0.173 (0.514)	0.113
CHIN_1	-0.00646 (0.00917)	0.496				
WCTA_1	-4.716*** (1.062)	19.712	-1.237* (0.669)	3.416		
CLCA_1	-0.0640 (0.0655)	0.954			0.370*** (0.0918)	16.251
FUTL_1	0.329 (0.238)	1.912				
INTWO_1	0.654* (0.343)	3.637				
RETA_1			0.0424 (0.0573)	0.548		
MVETL_1			-3.753*** (0.951)	15.568		
TLTA_1	-2.137** (0.907)	5.545			-0.138 (0.196)	0.496
OENEG_1	2.095*** (0.616)	11.558				
OhlsonSize_1	0.400*** (0.146)	7.527				
RelativeSize_1						
Price_1						
ExcessReturn_1						
SDReturn_1						
FirmAge_1						
Log Pseudolikelihood	-275.1780		-174.7317		-347.1652	
Pseudo R-Squared	0.2422		0.3914		0.0439	
Adjusted R-Squared	0.1790		0.3710		0.0440	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	Ratings		Shumway	
	Estimate	Wald-Chi Square	Estimate	Wald-Chi Square
Intercept	0.784	35.66	-0.215	66.82
Mean_1	-0.00630 (0.0391)	0.026		
EBITTA_1	0.848 (0.841)	1.016		
Sales_1				
NITA_1			-0.650 (0.634)	1.054
CHIN_1	-0.224*** (0.0798)	7.882		
WCTA_1	-5.332*** (1.476)	13.056		
CLCA_1				
FUTL_1				
INTWO_1				
RETA_1				
MVETL_1				
OENEG_1				
TLTA_1	0.578 (0.524)	1.217	-1.116 (0.800)	1.945
OhlsonSize_1				
RelativeSize_1			-0.669*** (0.202)	11.000
Price_1	-1.124*** (0.210)	28.552		
ExcessReturn_1	0.0625 (0.313)	0.040	-0.623 (0.742)	0.706
SDReturn_1	23.74*** (7.577)	9.814	14.08** (6.249)	5.076
FirmAge_1			-0.237*** (0.0765)	9.596
Log Pseudolikelihood	-124.9652		-301.7125	
Pseudo R-Squared	0.2178		0.1691	
Adjusted R-Squared	0.2780		0.1470	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 3: ROC Curves for Bankruptcy Models (Ohlson, Altman, Zmijewski, Shumway, and Ratings, respectively from left to right)

Figure 3 plots the Receiver Operating Characteristic (ROC) generated from the various models: Ohlson's Altman's, Zmijewski's, Shumway's, and Ratings' models. The ROC for Ohlson's model is 0.8340, the ROC for Altman's model is 0.9084, the ROC for Zmijewski's model is 0.5391, the ROC for Shumway's model is 0.7920, and the ROC for Ratings' model is 0.8867.

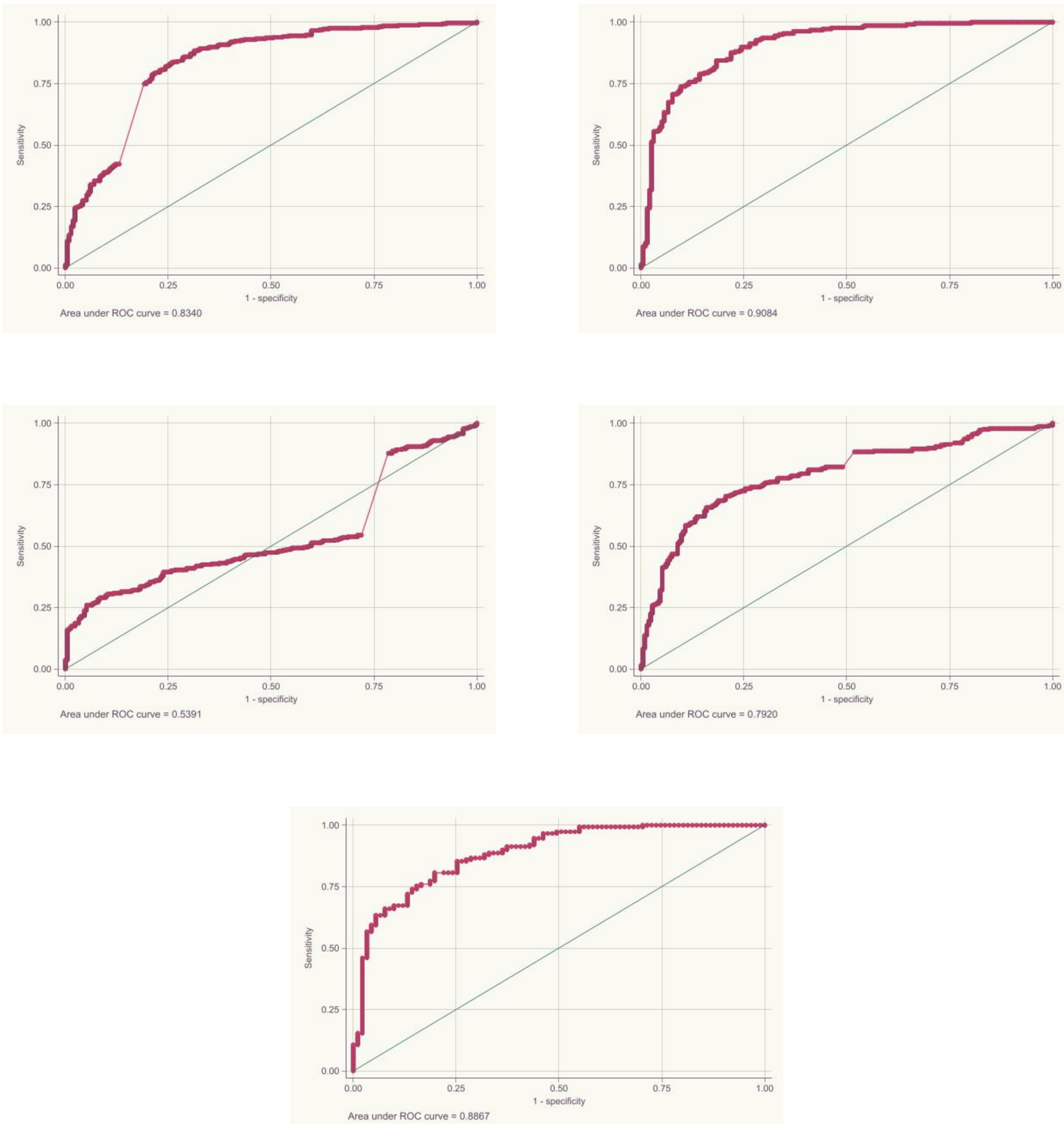


Figure 4: Optimal Cut-Off Points for Bankruptcy Models (Ohlson, Altman, Zmijewski, Shumway, and Ratings, respectively from left to right)

Figure 4 plots the optimal cut-off points generated from the various models: Ohlson's, Altman's, Zmijewski's, Shumway's, and Ratings' models. The horizontal line shows the optimal probability cutoff, while the vertical line shows the sensitivity/specificity of the models.

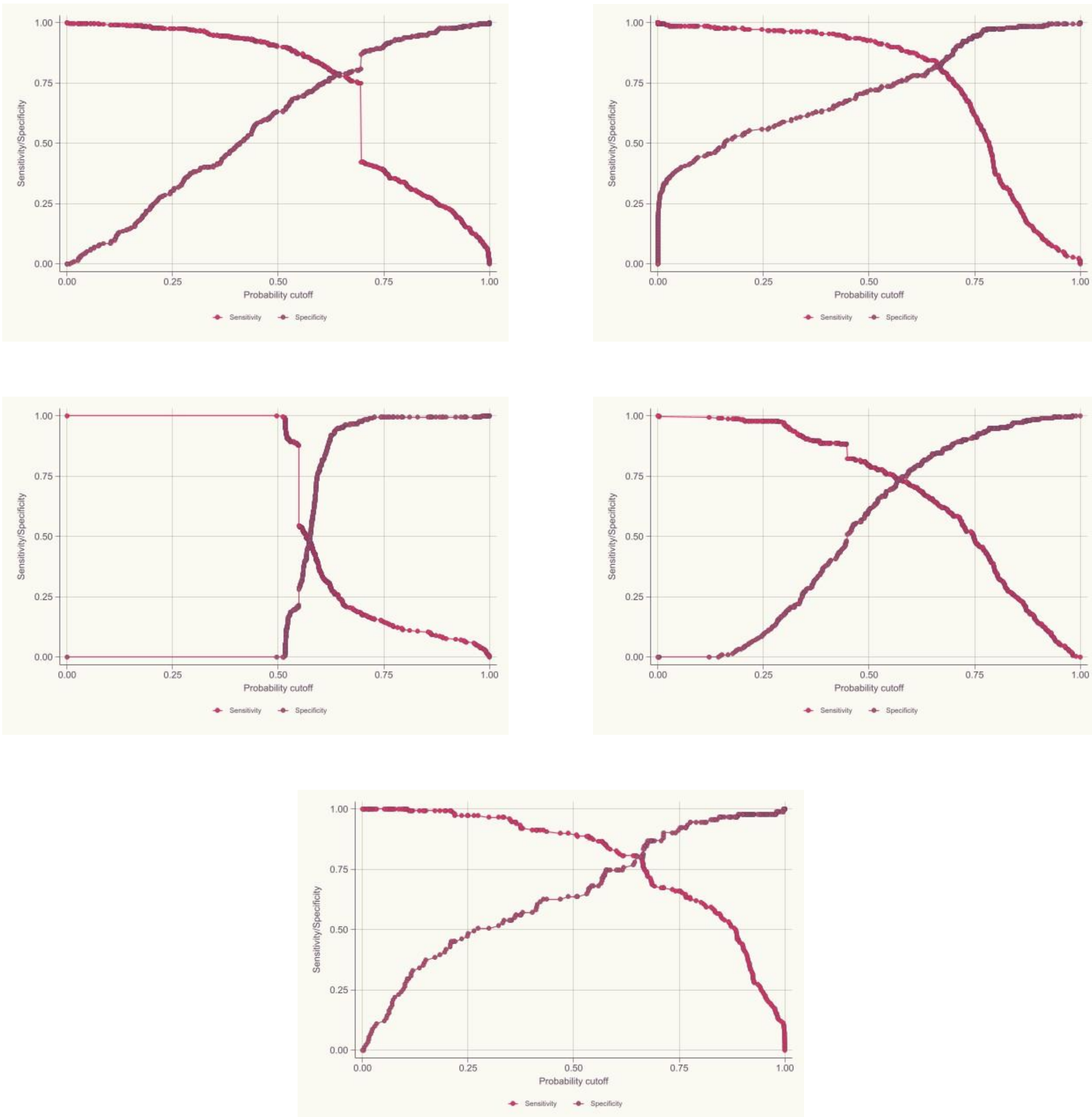


Table 3: Estimation Results for Bankruptcy Models with Ratings

Table 3 presents the parameter estimates from various bankruptcy prediction models with inclusion of mean employee satisfaction. The sample is from 2008 to 2020 and contains 327 bankruptcies and 327 non-bankrupt firms. Bold font signifies an estimate that is statistically significant. The adjusted Wald Chi-Square statistics from the logistic regression are presented adjacent to each parameter estimate. The Pseudo R-Squared for each model is reported. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Ohlson		**Altman**	
	Estimate	Wald-Chi Square	Estimate	Wald-Chi Square
Intercept	1.330	39.74	1.420	18.20
Mean_1	-0.203	0.385	-0.281	0.333
	(0.327)		(0.487)	
EBITTA_1			-4.241**	4.965
			(1.903)	
Sales_1			0.000606**	4.989
			(0.000271)	
NITA_1	-2.642**	4.828		
	(1.202)			
CHIN_1	-0.0365	0.366		
	(0.0604)			
WCTA_1	-5.757***	12.363	-1.765*	3.619
	(1.637)		(0.928)	
CLCA_1	0.0125	0.005		
	(0.173)			
FUTL_1	0.405	2.476		
	(0.258)			
INTWO_1	1.045**	4.275		
	(0.506)			
RETA_1			0.0483	0.304
			(0.0877)	
MVETL_1			-4.028***	7.968
			(1.427)	
OENEG_1	1.829**	4.530		
	(0.859)			
TLTA_1	-3.265***	14.777		
	(0.849)			
OhlsonSize_1	0.639***	14.761		
	(0.166)			
RelativeSize_1				
Price_1				
ExcessReturn_1				
SDReturn_1				
FirmAge_1				
Log Pseudolikelihood	-110.6051		-69.172861	
Pseudo R-Squared	0.3076		0.4696	
Adjusted R-Squared	0.2330		0.4160	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	Zmijewski		Shumway	
	Estimate	Wald-Chi Square	Estimate	Wald-Chi Square
Intercept	0.536	19.24	0.340	33.15
Mean_1	-0.121	0.157	-0.335	0.758
	(0.306)		(0.385)	
EBITTA_1				
Sales_1				
NITA_1	-0.691	0.847	-1.010	2.516
	(0.751)		(0.637)	
CHIN_1				
WCTA_1				
CLCA_1	0.644***	15.849		
	(0.162)			
FUTL_1				
INTWO_1				
RETA_1				
MVETL_1				
TLTA_1	-0.423	1.430	-1.732**	6.339
	(0.354)		(0.688)	
OENEG_1				
OhlsonSize_1				
RelativeSize_1			-0.852***	13.302
			(0.234)	
Price_1				
ExcessReturn_1			-0.0352	0.054
			(0.151)	
SDReturn_1			25.09***	18.409
			(5.848)	
FirmAge_1			-0.101	0.785
			(0.114)	
Log Pseudolikelihood	-148.2620		-129.5948	
Pseudo R-Squared	0.0719		0.1888	
Adjusted R-Squared	0.0410		0.1470	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 5: ROC Curves for Bankruptcy Models with Ratings (Ohlson, Altman, Zmijewski, and Shumway, respectively from left to right)

Figure 5 plots the Receiver Operating Characteristic (ROC) generated from the various models: Ohlson's Altman's, Zmijewski's, and Shumway's, but including the mean employee satisfaction in those models. The ROC for Ohlson's model is 0.8499, the ROC for Altman's model is 0.9182, the ROC for Zmijewski's model is 0.5987, and the ROC for Shumway's model is 0.7917.

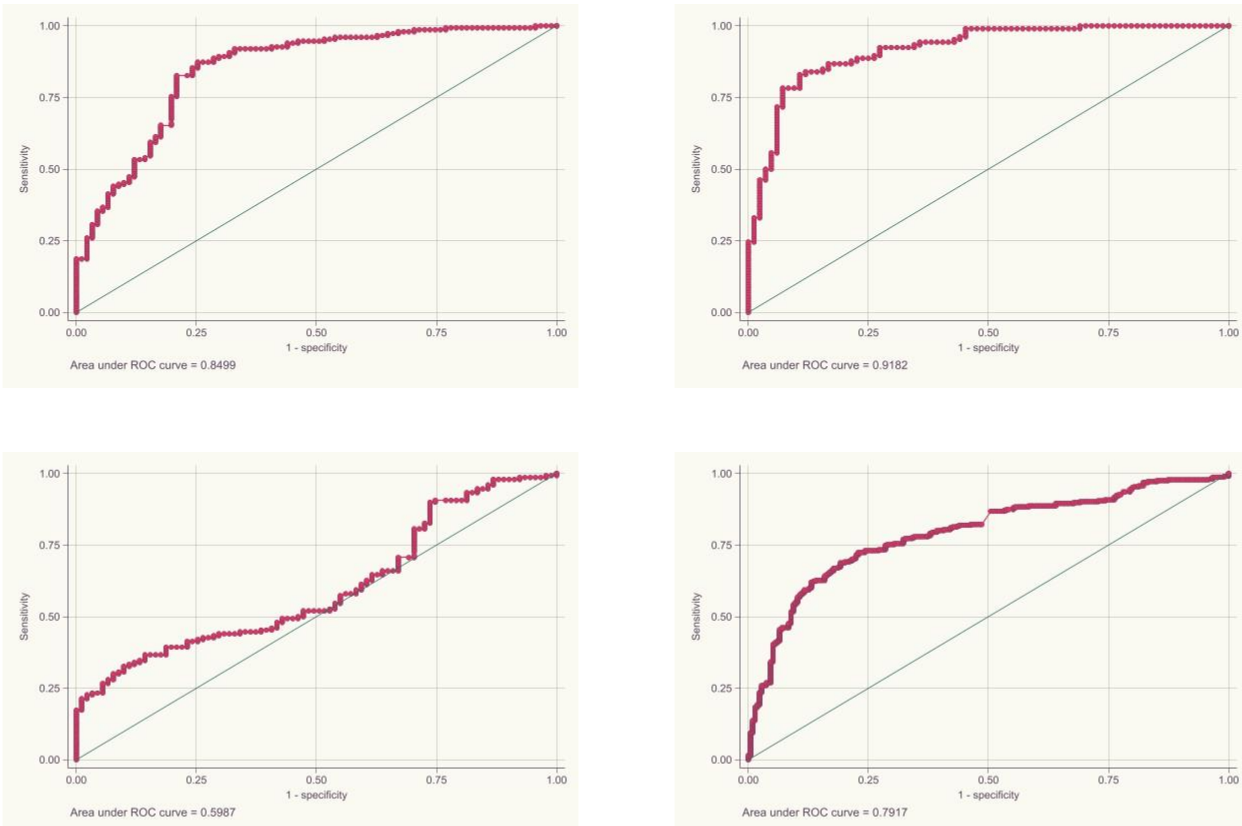


Figure 6: Optimal Cut-Off Points for Bankruptcy Models with Ratings (Ohlson, Altman, Zmijewski, and Shumway, respectively from left to right)

Figure 6 plots the optimal cut-off points generated from the various models: Ohlson's, Altman's, Zmijewski's, and Shumway's, but including the mean employee satisfaction in each model. The horizontal line shows the optimal probability cutoff, while the vertical line shows the sensitivity/specificity of the models.

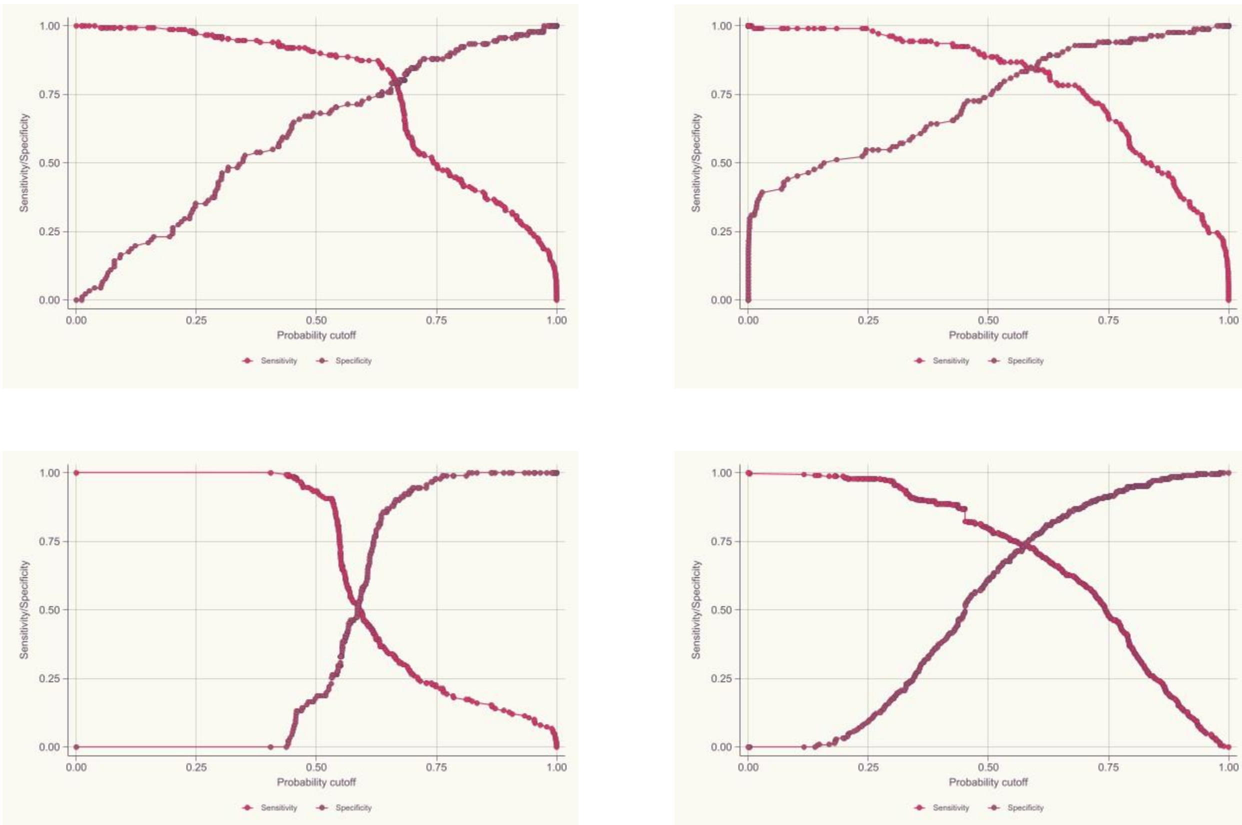


Table 4: Classification Rates for Bankruptcy Models

Table 4 presents the incidence of Type I errors (classifying a bankrupt firm as healthy) and Type II errors (classifying a healthy firm as bankrupt) and Total error (both Type I and Type II errors), according to model scores for Ohlson's, Altman's, Zmijewski's, Ratings', and Shumway's models.

	Ohlson			Altman			Zmijewski		
Model Score (Percentile)	Type I	Type II	Total	Type I	Type II	Total	Type I	Type II	Total
50	**0.577982**	**0.130841**	**0.708823**	**0.201835**	**0.167513**	**0.369348**	0.519878	0.528037	1.047915
70	0.584098	0.121495	0.705593	0.458716	0.030457	0.489172	0.645260	0.214953	0.860213
71	0.596330	0.112150	0.708480	0.477064	0.030457	0.507521	0.654434	0.200935	0.855369
72	0.605505	0.102804	0.708308	0.490826	0.025381	0.516206	0.663609	0.191589	0.855197
73	0.614679	0.093458	0.708137	0.509174	0.025381	0.534555	0.672783	0.182243	0.855026
74	0.626911	0.084112	0.711024	0.532110	0.025381	0.557491	0.678899	0.163551	0.842451
75	0.642202	0.084112	0.726314	0.550459	0.025381	0.575840	0.685015	0.149533	0.834548
76	0.651376	0.070094	0.721470	0.568807	0.025381	0.594188	0.691132	0.130841	0.821973
77	0.660551	0.060748	0.721298	0.587156	0.025381	0.612537	0.694190	0.112150	0.806339
78	0.675841	0.060748	0.736589	0.605505	0.025381	0.630885	0.700306	0.098131	0.798437
79	0.691132	0.056075	0.747206	0.623853	0.025381	0.649234	0.709480	0.084112	0.793592
80	0.703364	0.051402	0.754766	0.642202	0.025381	0.667583	0.721713	0.079439	0.801152
81	0.721713	0.051402	0.773114	0.665138	0.025381	0.690518	0.733945	0.070094	0.804038
82	0.730887	0.042056	0.772943	0.683486	0.025381	0.708867	0.740061	0.056075	0.796136
83	0.746177	0.037383	0.783561	0.697248	0.020305	0.717552	0.755352	0.051402	0.806754
84	0.755352	0.028037	0.783389	0.715596	0.020305	0.735901	0.767584	0.046729	0.814313
85	0.767584	0.023365	0.790949	0.733945	0.020305	0.754250	0.782875	0.046729	0.829604
86	0.785933	0.023365	0.809297	0.752294	0.020305	0.772598	0.795107	0.037383	0.832490
87	0.801223	0.023365	0.824588	0.770642	0.015228	0.785871	0.807339	0.032710	0.840050
88	0.816514	0.018692	0.835205	0.788991	0.015228	0.804219	0.819572	0.023365	0.842936
89	0.831804	0.018692	0.850496	0.807339	0.015228	0.822568	0.828746	0.014019	0.842765
90	0.844037	0.014019	0.858055	0.825688	0.015228	0.840917	0.840979	0.009346	0.850324
91	0.862385	0.014019	0.876404	0.844037	0.015228	0.859265	0.856269	0.004673	0.860942
92	0.874618	0.009346	0.883964	0.862385	0.015228	0.877614	0.871560	0.004673	0.876233
93	0.889908	0.004673	0.894581	0.880734	0.015228	0.895962	0.889908	0.004673	0.894581
94	0.905199	0.004673	0.909872	0.899083	0.010152	0.909235	0.905199	0.004673	0.909872
95	0.920489	0.004673	0.925162	0.912844	0.005076	0.917920	0.920489	0.004673	0.925162
96	0.938838	0.004673	0.943511	0.931193	0.005076	0.936269	0.938838	0.004673	0.943511
97	0.954128	0.004673	0.958801	0.949541	0.005076	0.954617	0.954128	0.004673	0.958801
98	0.972477	0.004673	0.977150	0.967890	0.005076	0.972966	0.969419	0	0.969419
99	0.987768	0.004673	0.992441	0.986239	0.005076	0.991315	0.984710	0	0.984710

	Ratings			Shumway		
Model Score (Percentile)	Type I	Type II	Total	Type I	Type II	Total
50	**0.3730887**	**0.1448598**	**0.5179485**	0.308869	0.205608	0.514476
70	0.5474006	0.0654206	0.6128212	0.550459	0.070094	0.620552
71	0.5626912	0.0607477	0.6234388	0.565749	0.065421	0.631170
72	0.5749236	0.0560748	0.6309983	0.577982	0.060748	0.638729
73	0.5871559	0.0514019	0.6385578	0.587156	0.051402	0.638558
74	0.6024465	0.0467290	0.6491755	0.605505	0.051402	0.656906
75	0.6177370	0.0467290	0.6644660	0.620795	0.051402	0.672197
76	0.6299694	0.0373832	0.6673526	0.639144	0.051402	0.690546
77	0.6452599	0.0373832	0.6826431	0.654434	0.051402	0.705836
78	0.6605505	0.0373832	0.6979337	0.669725	0.051402	0.721127
79	0.6788991	0.0373832	0.7162823	0.685015	0.046729	0.731744
80	0.6941896	0.0373832	0.7315728	0.700306	0.046729	0.747035
81	0.7125382	0.0373832	0.7499214	0.718655	0.046729	0.765383
82	0.7278287	0.0373832	0.7652119	0.730887	0.042056	0.772943
83	0.7461774	0.0373832	0.7835606	0.740061	0.028037	0.768099
84	0.7614679	0.0373832	0.7988511	0.755352	0.028037	0.783389
85	0.7767584	0.0373832	0.8141416	0.770642	0.028037	0.798680
86	0.7951070	0.0373832	0.8324902	0.785933	0.023365	0.809297
87	0.8073394	0.0327103	0.8400497	0.801223	0.023365	0.824588
88	0.8226300	0.0280374	0.8506674	0.816514	0.018692	0.835205
89	0.8379205	0.0280374	0.8659579	0.828746	0.014019	0.842765
90	0.8501529	0.0233645	0.8735174	0.844037	0.014019	0.858055
91	0.8685015	0.0233645	0.8918660	0.862385	0.014019	0.876404
92	0.8807340	0.0186916	0.8994256	0.874618	0.009346	0.883964
93	0.8960245	0.0140187	0.9100432	0.892966	0.009346	0.902312
94	0.9082569	0.0093458	0.9176027	0.908257	0.009346	0.917603
95	0.9204893	0.0046729	0.9251622	0.920489	0.004673	0.925162
96	0.9388379	0.0046729	0.9435108	0.938838	0.004673	0.943511
97	0.9541284	0.0046729	0.9588013	0.954128	0.004673	0.958801
98	0.9724771	0.0046729	0.9771500	0.972477	0.004673	0.977150
99	0.9847095	0	0.9847095	0.984710	0	0.984710

Figure 7: Classification Rates for Bankruptcy Models (Ohlson, Altman, Zmijewski, Shumway, and Ratings, respectively from left to right)

Figure 7 presents classification rates for the various bankruptcy models – Ohlson's, Altman's, Zmijewski's, Shumway's, and Ratings' models – from the tests in Table 4. The horizontal axis plots the percentile model score, while the vertical axis plots the Total error (Type I and Type II errors) for each respective model.

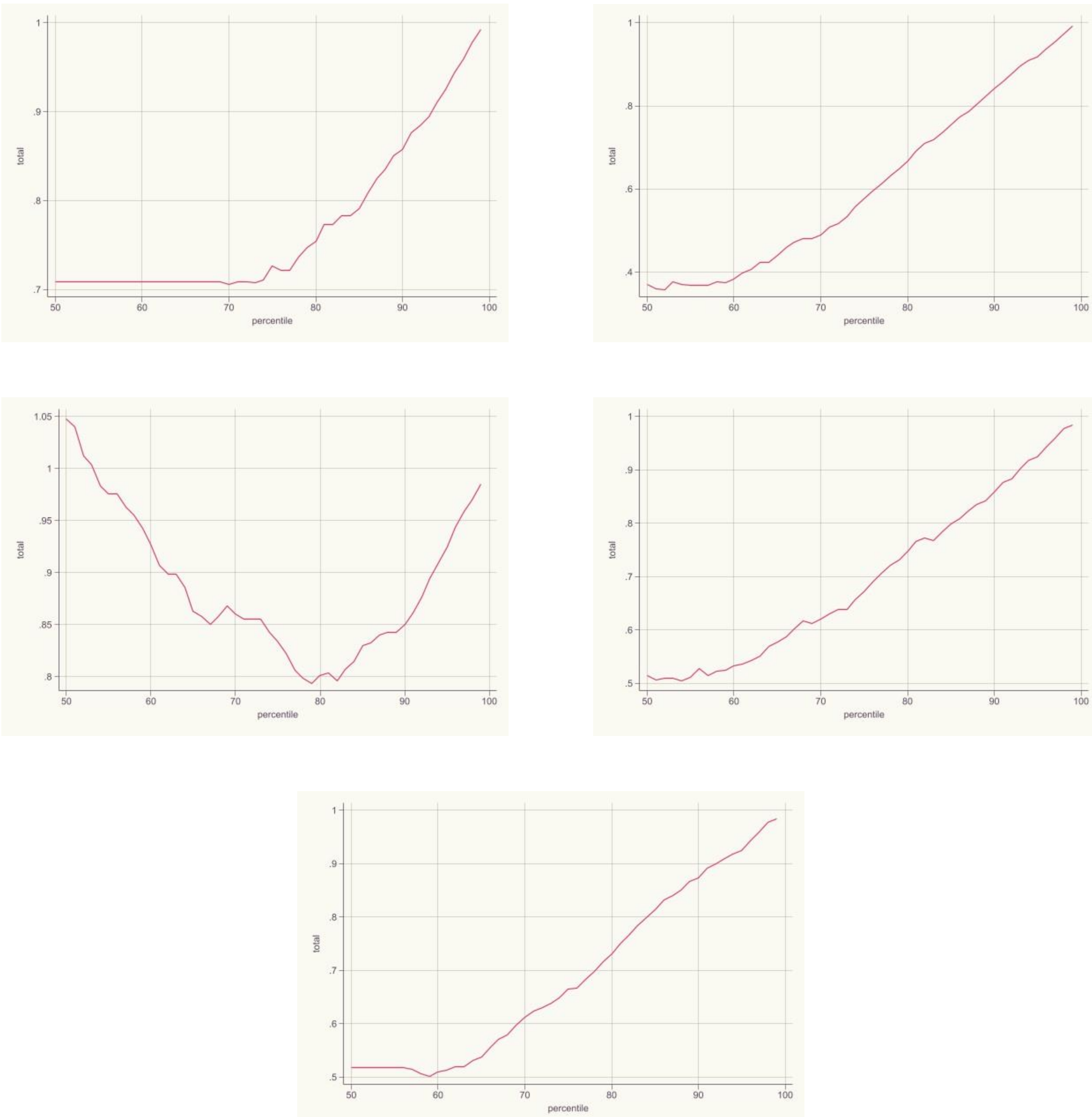


Table 5: Classification Rates for Bankruptcy Models with Ratings

Table 5 presents the incidence of Type I errors (classifying a bankrupt firm as healthy) and Type II errors (classifying a healthy firm as bankrupt) and Total error (both Type I and Type II errors), according to model scores for Ohlson's, Altman's, Zmijewski's, and Shumway's models with inclusion of mean of employee satisfaction.

	Ohlson			**Altman**		
Model Score (Percentile)	Type I	Type II	Total	Type I	Type II	Total
50	**0.299694**	**0.191589**	**0.491283**	**0.174312**	**0.137056**	**0.311368**
70	0.584098	0.121495	0.705593	0.477064	0.050761	0.527826
71	0.596330	0.112150	0.708480	0.495413	0.050761	0.546174
72	0.605505	0.102804	0.708308	0.509174	0.045685	0.554860
73	0.611621	0.088785	0.700406	0.527523	0.045685	0.573208
74	0.623853	0.079439	0.703293	0.550459	0.045685	0.596144
75	0.636086	0.074766	0.710852	0.568807	0.045685	0.614493
76	0.654434	0.074766	0.729201	0.587156	0.045685	0.632841
77	0.663609	0.065421	0.729029	0.605505	0.045685	0.651190
78	0.675841	0.060748	0.736589	0.623853	0.045685	0.669539
79	0.694190	0.060748	0.754937	0.637615	0.040609	0.678224
80	0.703364	0.051402	0.754766	0.655963	0.040609	0.696572
81	0.721713	0.051402	0.773114	0.674312	0.035533	0.709845
82	0.730887	0.042056	0.772943	0.692661	0.035533	0.728194
83	0.746177	0.037383	0.783561	0.706422	0.030457	0.736879
84	0.755352	0.028037	0.783389	0.724771	0.030457	0.755228
85	0.767584	0.023365	0.790949	0.743119	0.030457	0.773576
86	0.785933	0.023365	0.809297	0.756881	0.025381	0.782262
87	0.801223	0.023365	0.824588	0.779817	0.025381	0.805197
88	0.816514	0.018692	0.835205	0.798165	0.025381	0.823546
89	0.831804	0.018692	0.850496	0.816514	0.025381	0.841895
90	0.844037	0.014019	0.858055	0.834862	0.025381	0.860243
91	0.859327	0.009346	0.868673	0.848624	0.020305	0.868928
92	0.874618	0.009346	0.883964	0.866973	0.020305	0.887277
93	0.889908	0.004673	0.894581	0.885321	0.020305	0.905626
94	0.905199	0.004673	0.909872	0.908257	0.020305	0.928561
95	0.920489	0.004673	0.925162	0.926606	0.020305	0.946910
96	0.938838	0.004673	0.943511	0.944954	0.020305	0.965259
97	0.954128	0.004673	0.958801	0.958716	0.015228	0.973944
98	0.972477	0.004673	0.977150	0.972477	0.010152	0.982629
99	0.987768	0.004673	0.992441	1	0	1

	Zmijewski			Shumway		
Model Score (Percentile)	Type I	Type II	Total	Type I	Type II	Total
50	0.584098	0.467290	1.051388	0.333333	0.242991	0.576324
70	0.663609	0.242991	0.906599	0.571865	0.102804	0.674669
71	0.663609	0.214953	0.878562	0.581040	0.088785	0.669825
72	0.672783	0.205608	0.878390	0.593272	0.084112	0.677384
73	0.672783	0.182243	0.855026	0.602447	0.074766	0.677213
74	0.688073	0.177570	0.865644	0.620795	0.074766	0.695562
75	0.697248	0.168224	0.865472	0.633028	0.070094	0.703121
76	0.709480	0.158879	0.868359	0.645260	0.060748	0.706008
77	0.715596	0.144860	0.860456	0.660551	0.060748	0.721298
78	0.727829	0.140187	0.868016	0.672783	0.056075	0.728858
79	0.737003	0.126168	0.863171	0.685015	0.046729	0.731744
80	0.743119	0.112150	0.855269	0.697248	0.042056	0.739304
81	0.755352	0.102804	0.858155	0.715596	0.042056	0.757652
82	0.758410	0.084112	0.842522	0.727829	0.037383	0.765212
83	0.764526	0.065421	0.829947	0.743119	0.032710	0.775830
84	0.776758	0.060748	0.837506	0.755352	0.028037	0.783389
85	0.785933	0.051402	0.837335	0.770642	0.028037	0.798680
86	0.801223	0.046729	0.847952	0.788991	0.028037	0.817028
87	0.810398	0.037383	0.847781	0.804281	0.028037	0.832319
88	0.822630	0.028037	0.850667	0.819572	0.023365	0.842936
89	0.831804	0.018692	0.850496	0.828746	0.014019	0.842765
90	0.844037	0.014019	0.858055	0.840979	0.009346	0.850324
91	0.859327	0.009346	0.868673	0.856269	0.004673	0.860942
92	0.871560	0.004673	0.876233	0.871560	0.004673	0.876233
93	0.889908	0.004673	0.894581	0.889908	0.004673	0.894581
94	0.905199	0.004673	0.909872	0.905199	0.004673	0.909872
95	0.920489	0.004673	0.925162	0.920489	0.004673	0.925162
96	0.938838	0.004673	0.943511	0.938838	0.004673	0.943511
97	0.954128	0.004673	0.958801	0.954128	0.004673	0.958801
98	0.969419	0	0.969419	0.972477	0.004673	0.977150
99	0.984710	0	0.98471	0.987768	0.004673	0.992441

Figure 8: Classification Rates for Bankruptcy Models with Ratings (Ohlson, Altman, Zmijewski, and Shumway, respectively from left to right)

Figure 8 presents classification rates for the various bankruptcy models – Ohlson's, Altman's, Zmijewski's, and Shumway's models – from the tests in Table 5. The horizontal axis plots the percentile model score, while the vertical axis plots the Total error (Type I and Type II errors) for each respective model.

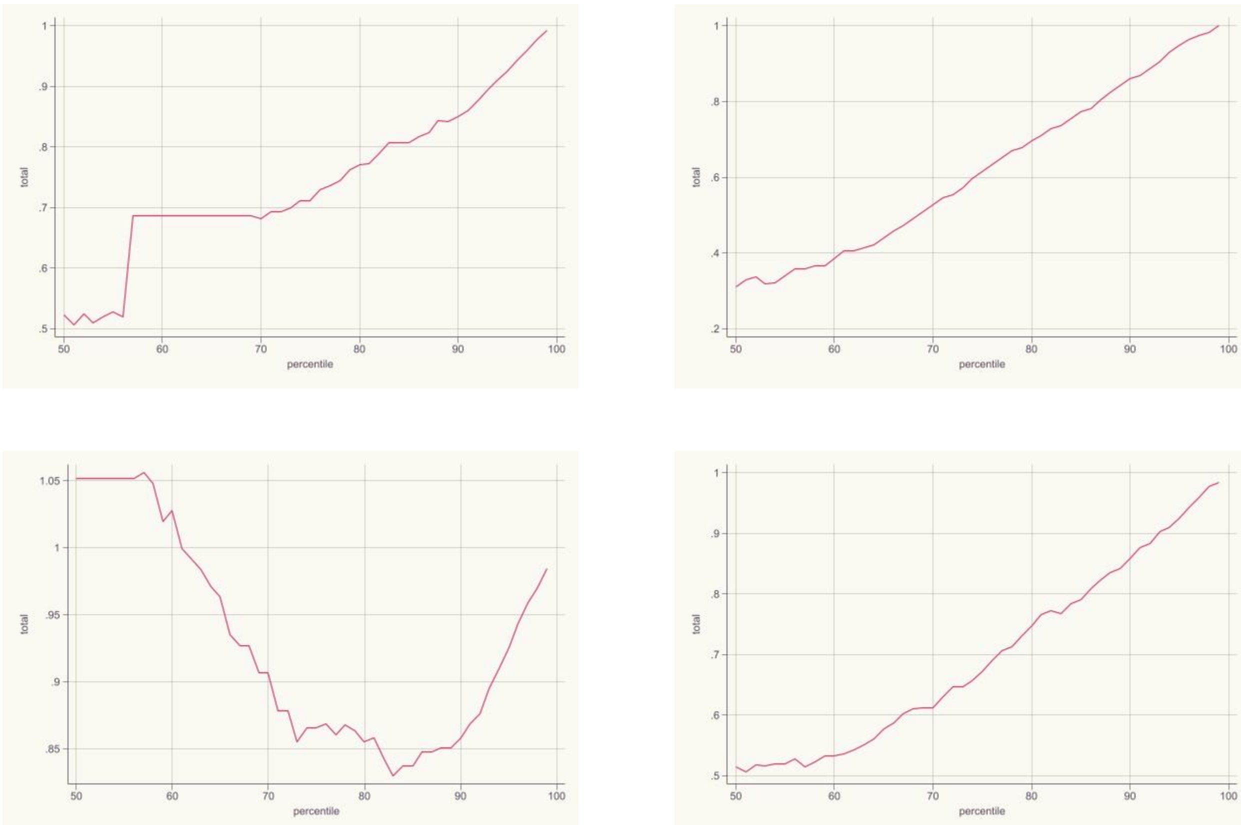


Table 6: Out-of-Sample Forecast Accuracy for Bankruptcy Models

Table 6 reports statistics relating to out-of-sample performance based on the rolling windows method. Observations are classified into deciles based on these bankruptcy probabilities (Decile 1 being the lowest). Actual and predicted probabilities are presented together with Wald-Chi Square statistic. The results for the following models are presented – Ohlson's, Altman's, Zmijewski's, Ratings', and Shumway's models.

	Ohlson			**Altman**		
Decile	Actual	Predicted	Chi-Square Stats	Actual	Predicted	Chi-Square Stats
1	5	2.939442	1.444458	1	0.00227	438.5925
2	13	11.30596	0.253828	2	0.646224	2.836032
3	15	18.87824	0.796724	5	7.177679	0.660699
4	22	28.22304	1.372149	13	17.99771	1.387795
5	119	95.85053	5.590977	23	26.08369	0.364563
6	-	-	-	23	28.72941	1.142599
7	15	15.09268	0.000569	33	31.52358	0.069149
8	42	44.13479	0.103259	40	32.35796	1.804834
9	44	50.51492	0.840230	40	35.31721	0.620902
10	52	53.65069	0.050787	38	38.16425	0.000707

	Zmijewski			**Ratings**		
Decile	Actual	Predicted	Chi-Square Stats	Actual	Predicted	Chi-Square Stats
1	29	28.43557	0.011204	5	2.124327	**3.892759**
2	120	83.91169	15.52068	7	7.020329	**0.0000589**
3	-	-	-	23	16.44145	**2.616228**
4	2	4.407028	1.314669	22	27.70281	**1.173961**
5	19	30.41686	4.285279	65	58.7226	**0.6710498**
6	17	31.31273	6.542204	11	14.56417	**0.8722303**
7	24	31.94927	1.977851	46	45.55432	**0.0043603**
8	25	33.09774	1.981206	48	49.36272	**0.0376199**
9	39	35.80448	0.285198	51	52.12202	**0.0241535**
10	52	47.66462	0.394329	49	53.81192	**0.4302869**

Shumway			
Decile	Actual	Predicted	Chi-Square Stats
1	14	12.85991	0.101074
2	23	18.86195	0.907830
3	21	24.50963	0.502558
4	16	24.57253	2.990669
5	27	30.98908	0.513496
6	35	35.56945	0.009117
7	44	39.78366	0.446856
8	49	42.84460	0.884336
9	47	46.26814	0.011577
10	51	50.74106	0.001322

Table 7: Out-of-Sample Forecast Accuracy for Bankruptcy Models with Ratings

Table 7 reports statistics relating to out-of-sample performance based on the rolling windows method. Observations are classified into deciles based on these bankruptcy probabilities (Decile 1 being the lowest). Actual and predicted probabilities are presented together with Wald-Chi Square statistic. The results for the following models are presented – Ohlson's, Altman's, Zmijewski's, and Shumway's models with the inclusion of mean of employee satisfaction.

		Ohlson			**Altman**	
Decile	Actual	Predicted	Chi-Square Stats	Actual	Predicted	Chi-Square Stats
1	5	6.060679	0.185629	1	0.000797	1252.523
2	13	14.78499	0.215503	2	0.496633	4.550865
3	13	22.86834	4.258467	6	7.084383	0.165983
4	26	29.6532	0.450065	9	17.34269	4.013246
5	41	35.35978	0.899668	20	25.60863	1.228366
6	92	73.92711	4.418264	29	30.59115	0.082761
7	1	1.406342	0.117407	37	34.33963	0.206104
8	39	41.60372	0.162951	39	36.56068	0.162751
9	46	48.33648	0.112941	39	40.21312	0.036597
10	51	52.99935	0.075424	36	40.90007	0.587057

		Zmijewski			**Shumway**	
Decile	Actual	Predicted	Chi-Square Stats	Actual	Predicted	Chi-Square Stats
1	24	27.1957	0.37552	20	15.20559	1.511703
2	39	29.32692	3.190536	10	24.02962	8.191154
3	33	30.27457	0.245353	28	29.15312	0.045611
4	24	32.43731	2.194640	29	32.06665	0.293276
5	71	55.41735	4.381645	22	36.14635	5.536362
6	9	12.70476	1.080324	37	40.35836	0.279462
7	17	35.53437	9.667337	41	44.36263	0.254883
8	26	37.63573	3.597387	41	48.00068	1.021017
9	33	41.4557	1.724704	47	50.19291	0.20311
10	51	51.34185	0.002276	52	52.63543	0.007671

Table 8: Classification Rates Using Altman's Z-Score and Ratings' Model One, Two, and Three Years Before

Table 8 reports classification rates for Ratings' and Altman's Z-Score model which shows sensitivity, specificity, positive predictive value, negative predictive value, false positives, and false negatives. The results are reported for each model one, two, and three years before merger announcement for Ratings' and Altman's Z-Score models.

8.1. Classification Rates for Ratings' Model

Classified + if predicted Pr(D)	>= .5			
True D defined as dealnodeal !=	0	One Year Before	Two Years Before	Three Years Before
Sensitivity	Pr(+ D)	86.54%	81.65%	89.91%
Specificity	Pr(-~D)	62.15%	68.69%	43.46%
Positive predictive value	Pr(D +)	77.75%	79.94%	70.84%
Negative predictive value	Pr(~D -)	75.14%	71.01%	73.81%
False positive rate	Pr(+~D)	37.85%	31.31%	56.54%
False negative rate	Pr(- D)	13.46%	18.35%	10.09%
False positive rate – classified	Pr(~D +)	22.25%	20.06%	29.16%
False negative rate – classified	Pr(D -)	24.86%	28.99%	26.19%
Correctly classified		76.89%	76.52%	71.53%

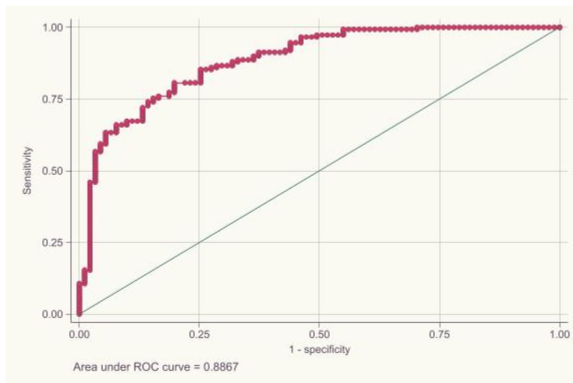
8.2. Classification Rates for Altman's Z-Score Model

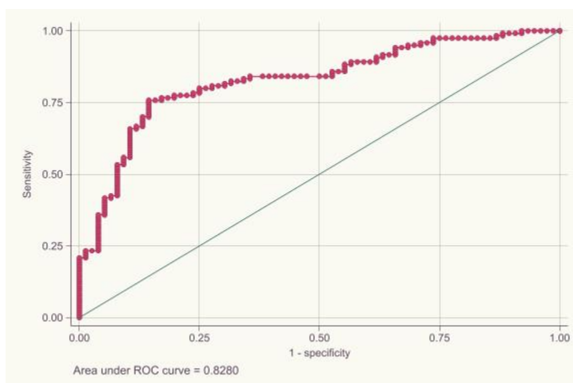
Classified + if predicted Pr(D)	>= .5			
True D defined as dealnodeal !=	0	One Year Before	Two Years Before	Three Years Before
Sensitivity	Pr(+ D)	98.67%	99.69%	99.08%
Specificity	Pr(-~D)	24.18%	1.87%	8.88%
Positive predictive value	Pr(D +)	68.20%	60.82%	62.43%
Negative predictive value	Pr(~D -)	91.67%	80.00%	86.36%
False positive rate	Pr(+~D)	75.82%	98.13%	91.12%
False negative rate	Pr(- D)	1.33%	0.31%	0.92%
False positive rate – classified	Pr(~D +)	31.80%	39.18%	37.57%
False negative rate – classified	Pr(D -)	8.33%	20.00%	13.64%
Correctly classified		70.54%	61.00%	63.40%

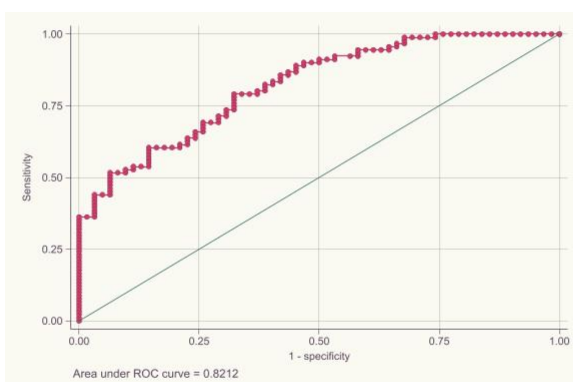
Figure 9: ROC Curves for Ratings' Model and Altman's Model One, Two, and Three Years Before, Respectively

Figure 9 plots the Receiver Operating Characteristic (ROC) generated from Ratings' and Ohlson's Altman's models. The ROC for Ratings' model one year before merger announcement is 0.8867, the ROC for Ratings' model two years before is 0.8280, the ROC for Ratings' model three years before is 0.8212, while the ROC for Altman's model one year before merger announcement is 0.9084, the ROC for Altman's model two years before merger announcement is 0.7808, and the ROC for Altman's model three years before is 0.7808.

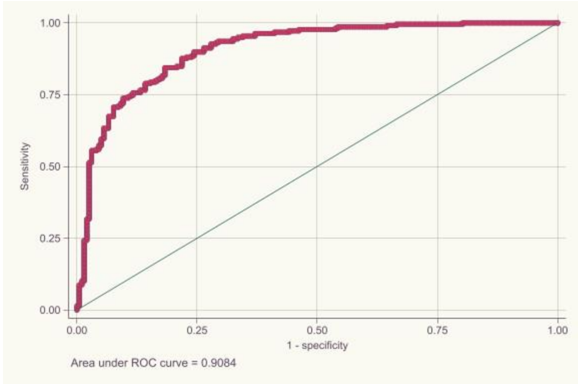
9.2. Ratings' Model

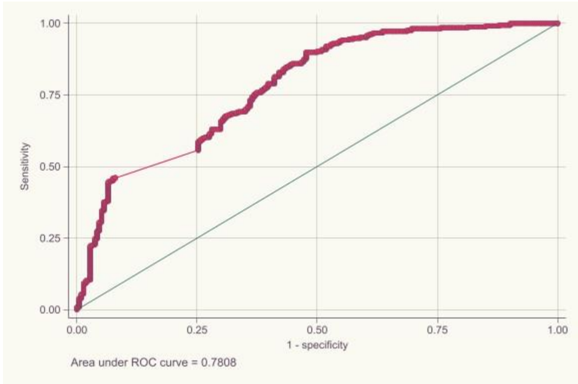






9.1. Altman's Model





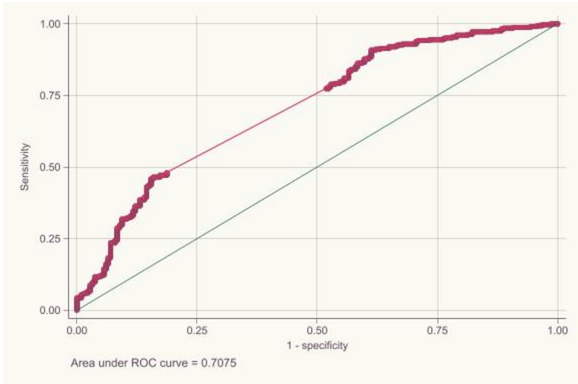


Table 9: Classification Rates Using Rating Categories' Model One, Two, and Three Years Before

Table 9 reports classification rates for breakdown of ratings' which shows sensitivity, specificity, positive predictive value, negative predictive value, false positives, and false negatives. The results are reported for each model one, two, and three years before merger announcement for Ratings' model using breakdown of rating categories.

Classified + if predicted Pr(D)	>= .5	Career_ Opps_1	Comp_ Benefits_1	Culture_ Values_1	Senior_ Lead_1	Worklife_ Balance_1	Overall_ Rating_1
True D defined as dealnodeal !=	0						
Sensitivity	Pr(+ D)	86.85%	86.85%	86.85%	86.85%	86.85%	86.85%
Specificity	Pr(-~D)	65.42%	65.42%	65.89%	66.36%	66.36%	66.36%
Positive predictive value	Pr(D +)	79.33%	79.33%	79.55%	79.78%	79.78%	79.78%
Negative predictive value	Pr(~D -)	76.50%	76.50%	76.63%	76.76%	76.76%	76.76%
False positive rate	Pr(+~D)	34.58%	34.58%	34.11%	33.64%	33.64%	33.64%
False negative rate	Pr(- D)	13.15%	13.15%	13.15%	13.15%	13.15%	13.15%
False positive rate – classified	Pr(~D +)	20.67%	20.67%	20.45%	20.22%	20.22%	20.22%
False negative rate – classified	Pr(D -)	23.50%	23.50%	23.37%	23.24%	23.24%	23.24%
Correctly classified		78.37%	78.37%	78.56%	78.74%	78.74%	78.74%

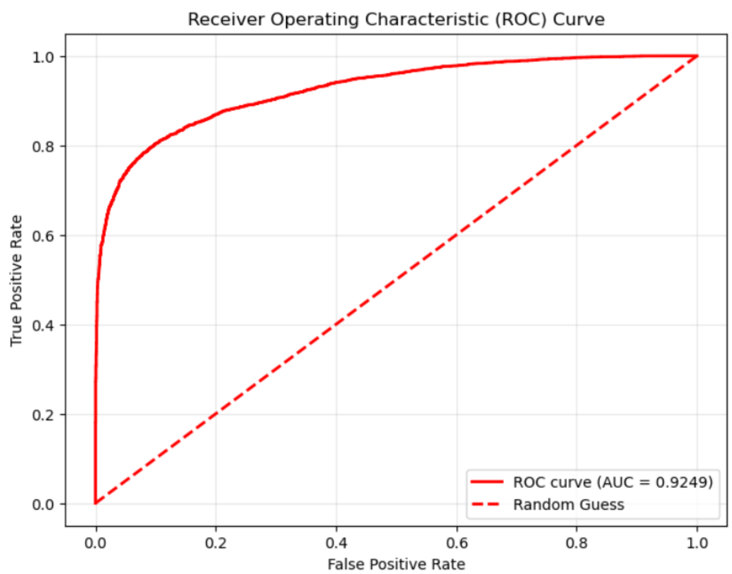
Classified + if predicted Pr(D)	>= .5	Career_ Opps_2	Comp_ Benefits_2	Culture_ Values_2	Senior_ Lead_2	Worklife_ Balance_2	Overall_ Rating_2
True D defined as dealnodeal !=	0						
Sensitivity	Pr(+ D)	81.96%	81.35%	81.65%	81.96%	81.96%	81.96%
Specificity	Pr(-~D)	62.15%	59.81%	61.68%	62.15%	61.21%	61.68%
Positive predictive value	Pr(D +)	76.79%	75.57%	76.50%	76.79%	76.35%	76.57%
Negative predictive value	Pr(~D -)	69.27%	67.72%	68.75%	69.27%	68.95%	69.11%
False positive rate	Pr(+~D)	37.85%	40.19%	38.32%	37.85%	38.79%	38.32%
False negative rate	Pr(- D)	18.04%	18.65%	18.35%	18.04%	18.04%	18.04%
False positive rate – classified	Pr(~D +)	23.21%	24.43%	23.50%	23.21%	23.65%	23.43%
False negative rate – classified	Pr(D -)	30.73%	32.28%	31.25%	30.73%	31.05%	30.89%
Correctly classified		74.12%	72.83%	73.75%	74.12%	73.75%	73.94%

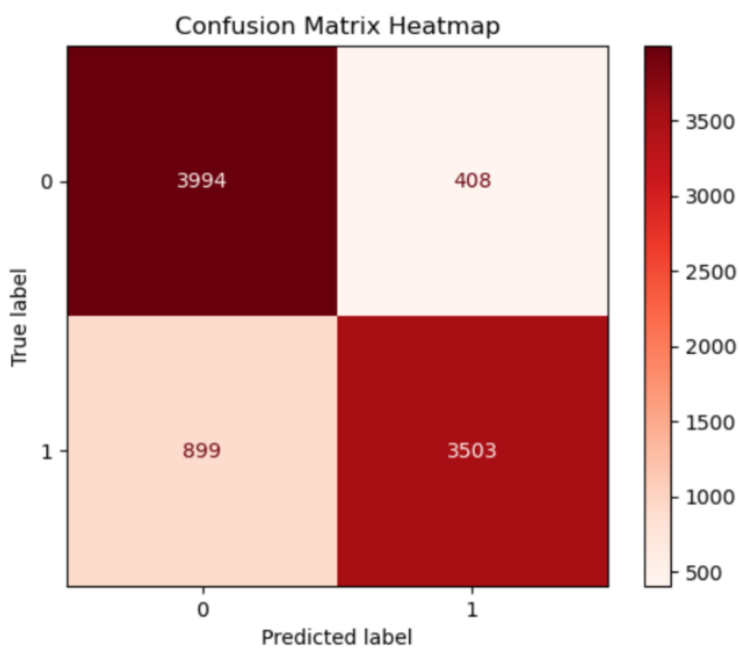
Classified + if predicted Pr(D)	>= .5	Career_ Opps_3	Comp_ Benefits_3	Culture_ Values_3	Senior_ Lead_3	Worklife_ Balance_3	Overall_ Rating_3
True D defined as dealnodeal !=	0						
Sensitivity	Pr(+ D)	81.96%	91.74%	89.91%	91.13%	92.05%	90.21%
Specificity	Pr(-~D)	61.68%	31.78%	35.51%	31.78%	32.24%	32.71%
Positive predictive value	Pr(D +)	76.57%	67.26%	68.06%	67.12%	67.49%	67.20%
Negative predictive value	Pr(~D -)	69.11%	71.58%	69.72%	70.10%	72.63%	68.63%
False positive rate	Pr(+~D)	38.32%	68.22%	64.49%	68.22%	67.76%	67.29%
False negative rate	Pr(- D)	18.04%	8.26%	10.09%	8.87%	7.95%	9.79%
False positive rate – classified	Pr(~D +)	23.43%	32.74%	31.94%	32.88%	32.51%	32.80%
False negative rate – classified	Pr(D -)	30.89%	28.42%	30.28%	29.90%	27.37%	31.37%
Correctly classified		73.94%	68.02%	68.39%	67.65%	68.39%	67.47%

Table 10 and Figure 10: Results from Machine Learning Models with Textual Reviews' Data

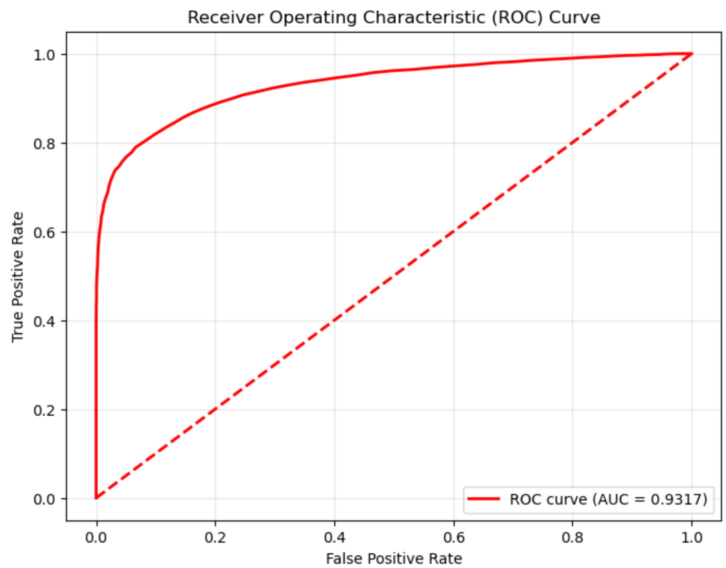
Table 10 and Figure 10 provide results of CatBoost model (10.1.), Random Forest model (10.2), Support Vector Machine model (10.3.), Regression model (10.4.), and Autoencoder Anomaly Detection model (10.5.) with employee satisfaction reviews. For the models, we use 6,344 actual failed company reviews and 22,009 pseudo company reviews. We use the pros, cons, and feedback sections in Glassdoor to make our bankruptcy filing predictions. Our test set contains 5,671 observations and our training set contains 22,682 observations. We use several models to test bankruptcy predictions – CatBoost model, random forest model, and support vector machine model. We calculate the confusion matrix for each model and compute the models' accuracy. We provide the results in the following figures and tables.

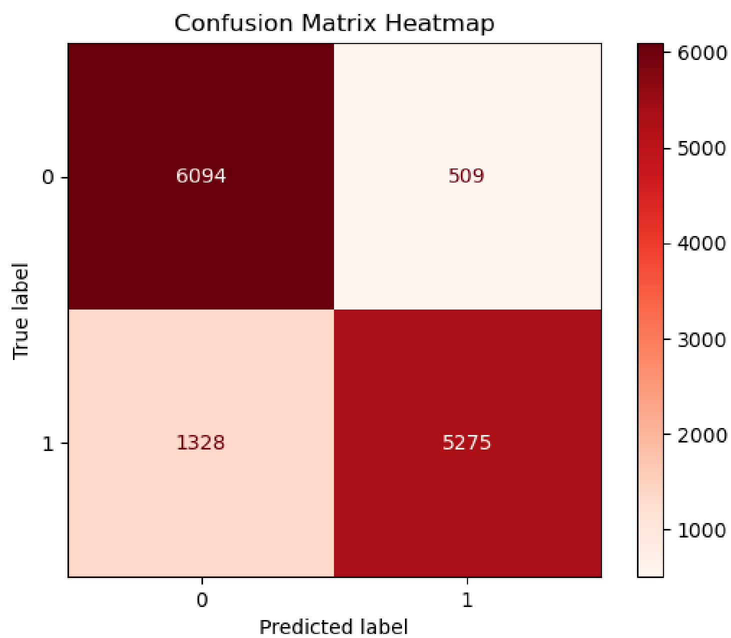
10.1. Catboost Model – ROC Curve and Error Rate Heat Map using Textual Reviews 1 Year Prior



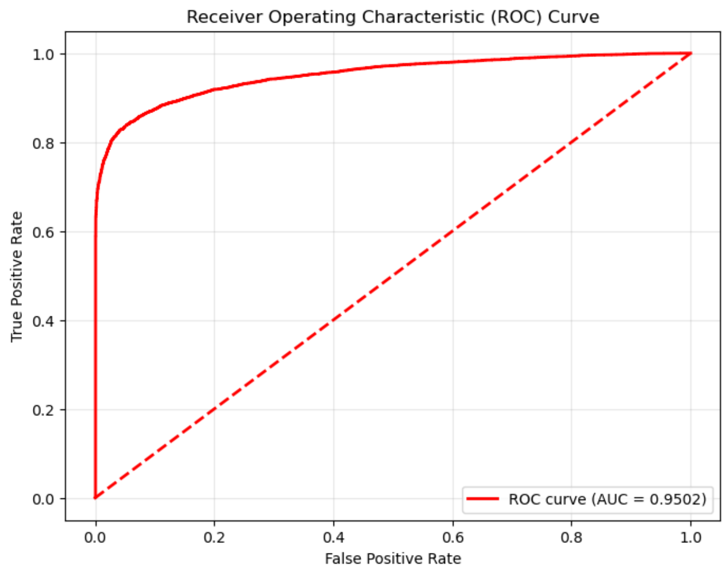


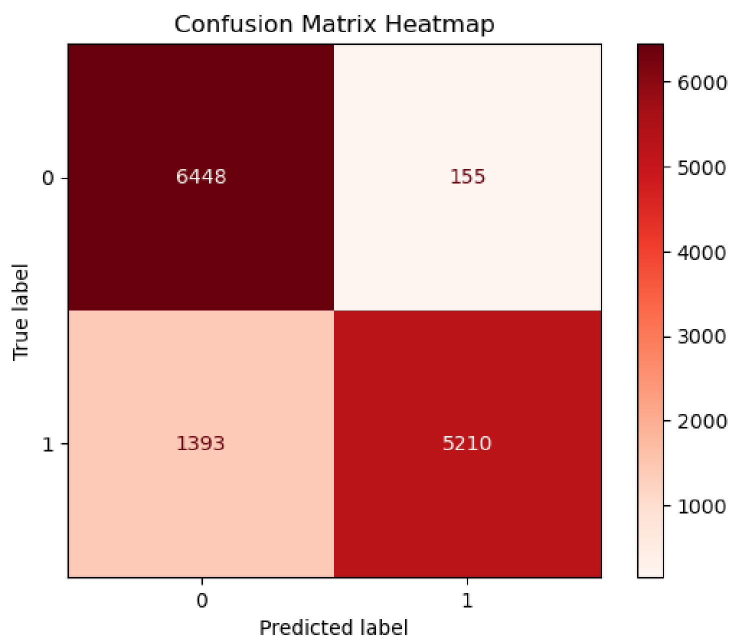
10.2. Random Forest Model – ROC Curve and Error Rate Heat Map using Textual Reviews 1 Year Prior



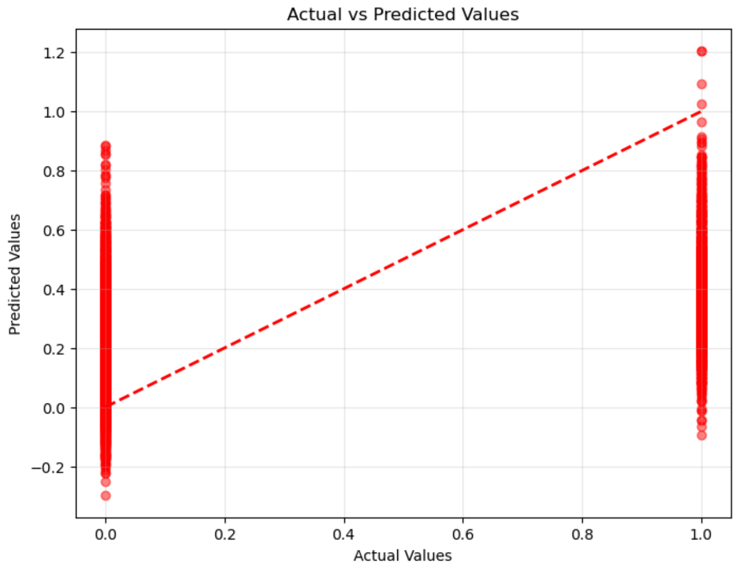


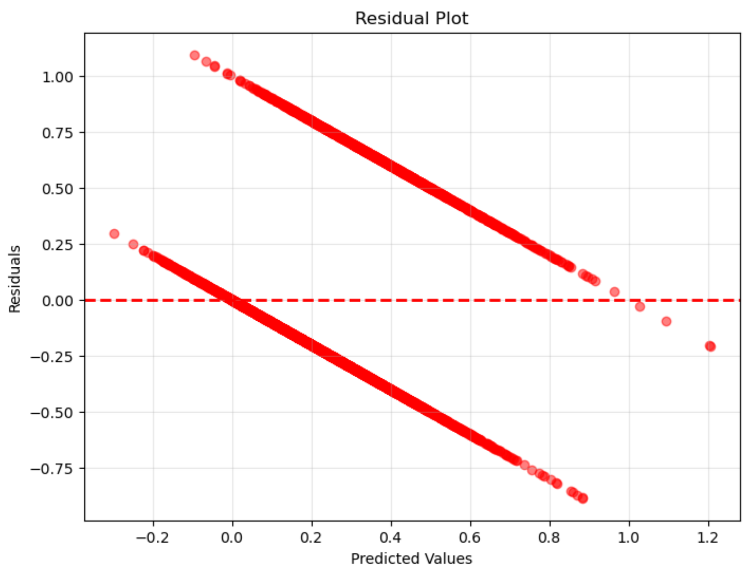
10.3. Support Vector Machine Model – ROC Curve and Error Rate Heat Map using Textual Reviews 1 Year Prior



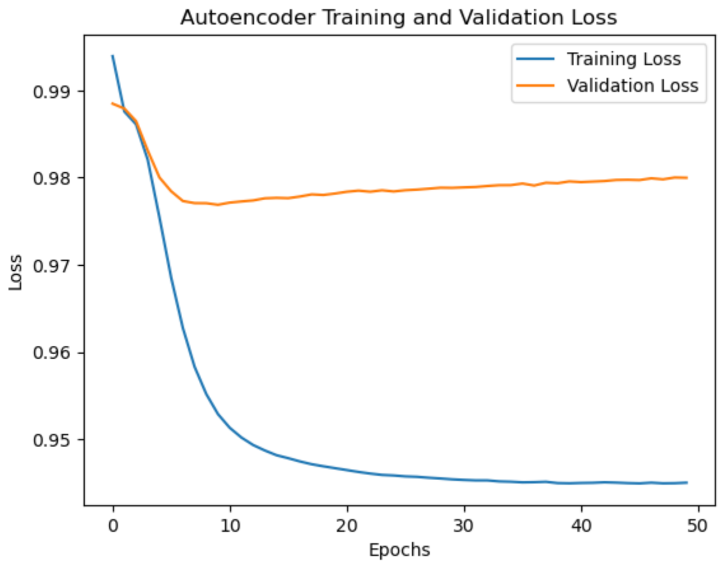


__10.4. Regression Model – Residual Plot and Actual vs. Predicted Plot using Textual Reviews 1 Year Prior__





10.5 Autoencoder Anomaly Detection – Training and Validation Loss and Reconstruction Error Distribution using Textual Reviews 1 Year Prior



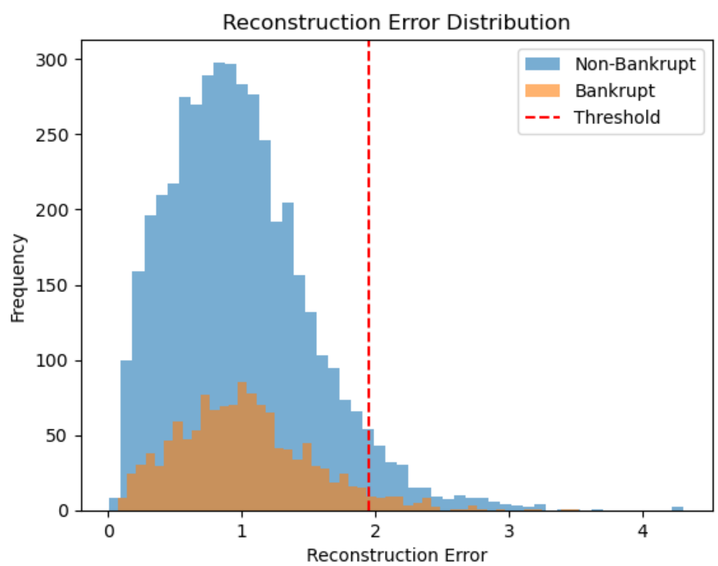


Table 11: Survival Analyses with Mean of Ratings and Standard Deviation of Ratings

Table 11 presents survival analyses with Cox regression for bankruptcy emergence with mean of employee satisfaction (Panel A) and standard deviation of employee satisfaction (Panel B) one, two, and three years before bankruptcy filings. Firm controls (EBITTA, WCTA, TLTA, CHIN, ExcessReturn, SDReturn, and Price) are also included in the model, as per in the employee satisfaction hazard model. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

11.1. Cox Regression for Bankruptcy Emergence with Mean of Employee Satisfaction

	Cox Regression for Bankruptcy Emergence with Mean					
Main Effects						
Mean_1	**0.451****	**(3.07)**				
Mean_2			**0.444****	**(2.98)**		
Mean_3					**0.494****	**(2.90)**
Controls						
EBITTA_1/2/3	0.000845	(0.00)	1.348	(1.67)	0.954	(1.78)
WCTA_1/2/3	-0.275	(-1.48)	-0.0471	(-0.09)	-1.040	(-1.86)
TLTA_1/2/3	0.00164	(0.02)	0.148	(0.75)	0.0170	(0.07)
CHIN_1/2/3	-0.00633	(-1.03)	0.0399*	(1.96)	-0.00401	(-0.78)
ExcessReturn_1/2/3	-0.243	(-1.17)	-0.0517	(-0.43)	-0.256	(-1.79)
SDReturn_1/2/3	-0.159	(-0.07)	2.336	(1.12)	0.600	(0.20)
Price_1/2/3	-0.0219	(-0.32)	0.0261	(0.38)	0.117	(1.69)
Observations	301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

11.2. Cox Regression for Bankruptcy Emergence with Standard Deviation of Employee Satisfaction

	Cox Regression for Bankruptcy Emergence with Standard Deviation					
Main Effects						
SD_1	**0.414****	**(2.66)**				
SD_2			**0.331***	**(2.11)**		
SD_3					**0.561****	**(3.10)**
Controls						
EBITTA_1/2/3	-0.0263	(-0.12)	1.355	(1.67)	0.885	(1.58)
WCTA_1/2/3	-0.304	(-1.65)	-0.0752	(-0.15)	-1.118*	(-1.99)
TLTA_1/2/3	-0.0101	(-0.10)	0.123	(0.63)	-0.0478	(-0.19)
CHIN_1/2/3	-0.00632	(-1.04)	0.0409*	(2.00)	-0.00357	(-0.67)
ExcessReturn_1/2/3	-0.222	(-1.09)	-0.0543	(-0.44)	-0.308*	(-2.13)
SDReturn_1/2/3	0.270	(0.12)	2.833	(1.42)	1.740	(0.59)
Price_1/2/3	-0.0221	(-0.33)	0.0301	(0.43)	0.136*	(1.98)
Observations	301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Table 12: Survival Analyses with Mean of Rating Categories

Table 12 presents survival analyses with Cox regression for bankruptcy emergence with mean of employee satisfaction rating categories (Career Opportunities, Compensation Benefits, Culture Values, Senior Leadership, Work-Life Balance, and Overall Rating) one, two, and three years before bankruptcy filings. Firm controls (EBITTA, WCTA, TLTA, CHIN, ExcessReturn, SDReturn, and Price) are also included in the model, as per in the employee satisfaction hazard model. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cox Regression for Bankruptcy Emergence with Mean Rating Categories							
Main Effects								
careeropps_mean_1	0.229	(1.42)						
compensation_mean_1			**0.352***	**(2.20)**				
culture_mean_1					**0.378***	**(2.34)**		
seniorleadership_mean_1							0.0832	(0.50)
Controls								
EBITTA_1	0.0368	(0.17)	0.0890	(0.40)	-0.0242	(-0.11)	0.0256	(0.12)
WCTA_1	-0.265	(-1.45)	-0.317	(-1.71)	-0.254	(-1.39)	-0.271	(-1.47)
TLTA_1	0.0159	(0.16)	0.00758	(0.08)	0.00826	(0.08)	0.0106	(0.11)
CHIN_1	-0.00646	(-1.05)	-0.00816	(-1.35)	-0.00617	(-1.01)	-0.00748	(-1.23)
ExcessReturn_1	-0.338	(-1.53)	-0.298	(-1.39)	-0.236	(-1.11)	-0.352	(-1.59)
SDReturn_1	-0.394	(-0.17)	-0.0457	(-0.02)	-0.332	(-0.15)	-0.296	(-0.13)
Price_1	-0.0127	(-0.19)	-0.00913	(-0.14)	-0.0183	(-0.27)	-0.00357	(-0.05)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergence with Mean Rating Categories			
Main Effects				
worklife_mean_1	0.176	(1.09)		
overall_mean_1			0.284	(1.62)
Controls				
EBITTA_1	0.0347	(0.16)	0.0170	(0.08)
WCTA_1	-0.273	(-1.49)	-0.260	(-1.42)
TLTA_1	0.00887	(0.09)	0.0126	(0.13)
CHIN_1	-0.00659	(-1.07)	-0.00619	(-1.01)
ExcessReturn_1	-0.347	(-1.58)	-0.333	(-1.50)
SDReturn_1	-0.338	(-0.15)	-0.505	(-0.22)
Price_1	-0.00249	(-0.04)	-0.0115	(-0.17)
Observations	301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergence with Mean Rating Categories							
Main Effects								
careeropps_mean_2	**0.727*** **	**(3.93)**						
compensation_mean_2			**0.802*** **	**(4.17)**				
culture_mean_2					**0.506** **	**(3.02)**		
seniorleadership_mean_2							**0.764*** **	**(4.17)**
Controls								
EBITTA_2	1.460	(1.80)	1.586*	(2.01)	1.328	(1.64)	1.533	(1.92)
WCTA_2	-0.131	(-0.25)	0.118	(0.23)	-0.127	(-0.25)	0.00776	(0.01)
TLTA_2	0.138	(0.71)	0.265	(1.36)	0.145	(0.74)	0.169	(0.87)
CHIN_2	0.0310	(1.56)	0.0472*	(2.17)	0.0376	(1.87)	0.0311	(1.59)
ExcessReturn_2	-0.0252	(-0.20)	-0.0427	(-0.36)	-0.0324	(-0.26)	-0.0697	(-0.60)
SDReturn_2	2.655	(1.33)	1.896	(0.93)	2.952	(1.48)	2.578	(1.29)
Price_2	0.0303	(0.45)	-0.0203	(-0.28)	0.0273	(0.40)	0.0212	(0.31)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergence with Mean Rating Categories			
Main Effects				
worklife_mean_2	**0.692*** **	**(3.75)**		
overall_mean_2			**0.824*** **	**(4.34)**
Controls				
EBITTA_2	1.505	(1.88)	1.483	(1.87)
WCTA_2	0.0142	(0.03)	-0.0468	(-0.09)
TLTA_2	0.183	(0.94)	0.226	(1.17)
CHIN_2	0.0443*	(2.04)	0.0298	(1.50)
ExcessReturn_2	-0.0962	(-0.82)	-0.00461	(-0.04)
SDReturn_2	2.280	(1.11)	2.994	(1.53)
Price_2	0.0284	(0.41)	0.00161	(0.02)
Observations	301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergency with Mean Rating Categories							
Main Effects								
careeropps_mean_3	**0.598****	**(3.04)**						
compensation_mean_3			**0.512****	**(2.63)**				
culture_mean_3					**0.430***	**(2.29)**		
seniorleadership_mean_3							**0.438***	**(2.23)**
Controls								
EBITTA_3	0.891	(1.64)	1.001*	(1.96)	1.041	(1.94)	1.028	(1.95)
WCTA_3	-1.050	(-1.90)	-0.958	(-1.68)	-0.999	(-1.79)	-0.904	(-1.62)
TLTA_3	0.0733	(0.30)	0.181	(0.74)	0.0615	(0.25)	0.132	(0.54)
CHIN_3	-0.00394	(-0.79)	-0.00352	(-0.69)	-0.00367	(-0.73)	-0.00335	(-0.66)
ExcessReturn_3	-0.318*	(-2.20)	-0.247	(-1.73)	-0.277	(-1.92)	-0.277	(-1.96)
SDReturn_3	1.071	(0.36)	0.700	(0.24)	2.347	(0.81)	0.969	(0.33)
Price_3	0.124	(1.82)	0.104	(1.49)	0.119	(1.73)	0.117	(1.69)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergence with Mean Rating Categories			
Main Effects				
worklife_mean_3	0.359	(1.75)		
overall_mean_3			**0.596****	**(3.11)**
Controls				
EBITTA_3	0.965	(1.87)	0.938	(1.86)
WCTA_3	-0.924	(-1.65)	-1.014	(-1.83)
TLTA_3	0.126	(0.51)	0.155	(0.63)
CHIN_3	-0.00307	(-0.59)	-0.00391	(-0.78)
ExcessReturn_3	-0.282	(-1.93)	-0.240	(-1.69)
SDReturn_3	1.332	(0.46)	0.214	(0.07)
Price_3	0.128	(1.86)	0.115	(1.66)
Observations	301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Online Appendix

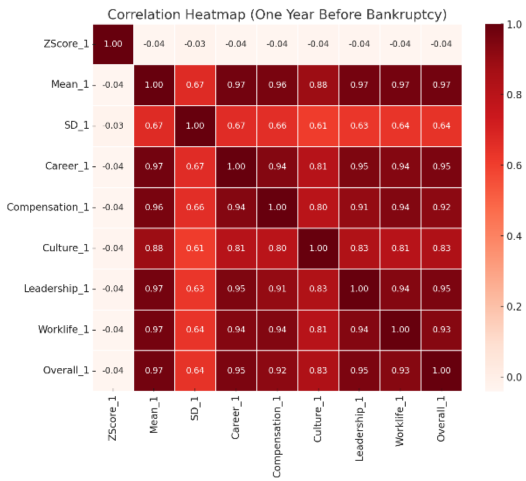
for

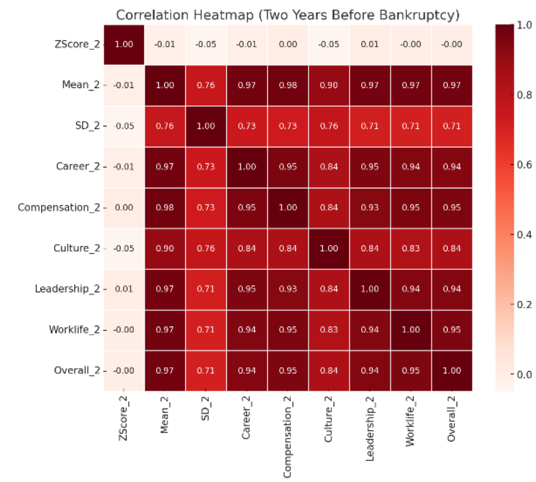
Predicting Bankruptcy: Ask the Employees

By John Knopf and Kristina Lalova

Figure A1: Heat Map of Correlation Matrix

Figure A1 presents correlation matrices between Z-Score and the mean of various rating categories for one, two, and three years before bankruptcy filings (from left to right, respectively). The figures display heatmaps of these correlations using a red scale, where darker shades represent stronger correlations. The matrices include correlations between Z-Score and aggregated sum of rating categories and between Z-Score and categories, such as Career Opportunities, Compensation and Benefits, Culture Values, Senior Leadership, Work-Life Balance, and Overall Rating.





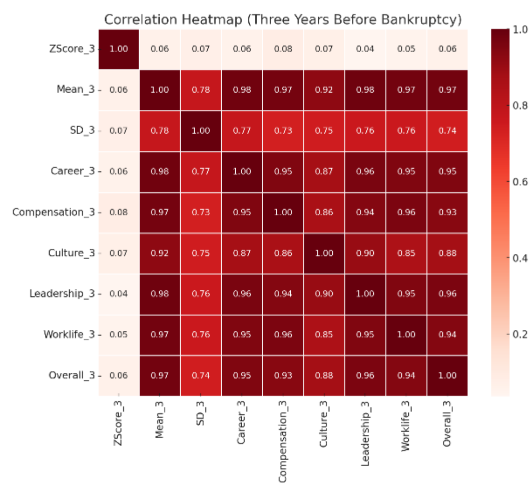


Table A2: Linear Discriminant Analysis with Unstandardized and Standardized Coefficients for Altman's Model (equivalent to Tables 2 and 3)

Table A2 presents Altman's model results equivalent to the results for Altman's model in a multi-period logistic setting. The discriminant analysis coefficients (both standardized and unstandardized) with the original Altman's model variables are presented in columns (1) and (2), while the discriminant analysis coefficients (both standardized and unstandardized) with the original Altman's model variables along with the employee satisfaction variable are presented in columns (3) and (4).

Predictors	**Unstandardized Coefficients (1)**	**Standardized Coefficients (2)**	**Unstandardized Coefficients (3)**	**Standardized Coefficients (4)**
Mean_1	-	-	0.00466	0.042905
WCTA_1	1.476932	0.550663	-1.4739	-0.54953
RETA_t1	-0.07409	-0.13922	0.072564	0.136357
EBITTA_t1	-0.16224	-0.0527	0.171106	0.055577
MVETL_t1	0.204272	0.501703	-0.20474	-0.50285
STA_t1	0.698869	0.60159	-0.69903	-0.60173
_cons	-0.72771	-	0.692398	-

Table A3: Error Rates for Altman's and Ratings' Models in a Discriminant Analysis Setting with Variable Information One, Two, and Three Years Before Bankruptcy Filing

Table A3 presents error rates, but in a discriminant analysis setting, for Altman's model and employee satisfaction model. We present stratified and unstratified error rates for both models one, two, and three years before bankruptcy filings with a prior of 0.5.

Employees' Model	**Bankruptcy?**								
	One Year Before			Two Years Before			Three Years Before		
Error rate	0	1	Total	0	1	Total	0	1	Total
Stratified	0.205768	0.268095	0.236931	0.315487	0.279221	0.297354	0.292589	0.41908	0.355835
Unstratified	0.291163	0.183687	0.237425	0.375824	0.220211	0.298017	0.331577	0.380586	0.356081
Priors	0.5	0.5		0.5	0.5		0.5	0.5	

Altman's Model	**Bankruptcy?**								
	One Year Before			Two Years Before			Three Years Before		
Error rate	0	1	Total	0	1	Total	0	1	Total
Stratified	0.371231	0.152863	0.262047	0.464176	0.226945	0.34556	0.536242	0.225025	0.380634
Unstratified	0.445255	0.082962	0.264108	0.513077	0.187525	0.350301	0.57001	0.198978	0.384494
Priors	0.5	0.5		0.5	0.5		0.5	0.5	

Table A4: Cox Regression for Bankruptcy Emergence with Z-Score

Table A4 presents survival analyses with Cox regression for bankruptcy emergence with Z-Score one, two, and three years before bankruptcy filings. Firm controls (WCTA, RETA, EBITTA, MVETL, and STA) are also included in the model, as per in the employee satisfaction hazard model. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cox Regression for Bankruptcy Emergence with Z-Score					
Main Effects						
Z-Score_1	-0.249	(-0.75)				
Z-Score_2			0.115	(0.53)		
Z-Score_3					**0.605***	**(3.23)**
Controls						
WCTA_1/2/3	-0.317	(-1.78)	-0.486	(-0.89)	-1.747**	(-2.72)
RETA_1/2/3	0.0532	(0.70)	-0.0936	(-0.99)	-0.142**	(-2.73)
EBITTA_1/2/3	-0.0493	(-0.22)	1.376	(1.76)	0.752	(1.54)
MVETL_1/2/3	0.177	(0.47)	0.0824	(0.87)	0.0555	(1.71)
STA_1/2/3	-0.00118	(-0.01)	-0.00910	(-0.07)	-0.0795	(-0.76)
Observations	301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Table A5: Survival Analyses with Standard Deviation of Rating Categories

Table A5 presents survival analyses with Cox regression for bankruptcy emergence with standard deviation of employee satisfaction rating categories (Career Opportunities, Compensation Benefits, Culture Values, Senior Leadership, Work-Life Balance, and Overall Rating) one, two, and three years before bankruptcy filings. Firm controls (EBITTA, WCTA, TLTA, CHIN, ExcessReturn, SDReturn, and Price) are also included in the model, as per in the employee satisfaction hazard model. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cox Regression for Bankruptcy Emergence with SD Rating Categories							
Main Effects								
careeropps_sd_1	**0.338***	**(2.12)**						
compensation_sd_1			**0.433****	**(2.65)**				
culture_sd_1					**0.520****	**(3.02)**		
seniorleadership_sd_1							**0.369***	**(2.29)**
Controls								
EBITTA_1	-0.0198	(-0.09)	-0.0575	(-0.26)	-0.0554	(-0.25)	-0.0186	(-0.08)
WCTA_1	-0.285	(-1.55)	-0.283	(-1.54)	-0.263	(-1.43)	-0.298	(-1.61)
TLTA_1	-0.00263	(-0.03)	-0.0132	(-0.13)	0.00631	(0.07)	-0.00728	(-0.07)
CHIN_1	-0.00638	(-1.05)	-0.00616	(-1.01)	-0.00575	(-0.94)	-0.00634	(-1.04)
ExcessReturn_1	-0.266	(-1.26)	-0.213	(-1.03)	-0.187	(-0.93)	-0.228	(-1.09)
SDReturn_1	0.0615	(0.03)	0.209	(0.10)	0.107	(0.05)	0.420	(0.19)
Price_1	-0.0201	(-0.30)	-0.0292	(-0.43)	-0.0373	(-0.55)	-0.0158	(-0.24)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

	Cox Regression for Bankruptcy Emergence with SD Rating Categories			
Main Effects				
worklife_sd_1	**0.356***	**(2.15)**		
overall_sd_1			**0.346***	**(2.14)**
Controls				
EBITTA_1	-0.0334	(-0.15)	-0.0159	(-0.07)
WCTA_1	-0.284	(-1.54)	-0.298	(-1.62)
TLTA_1	-0.00581	(-0.06)	-0.00582	(-0.06)
CHIN_1	-0.00638	(-1.05)	-0.00638	(-1.04)
ExcessReturn_1	-0.236	(-1.11)	-0.239	(-1.12)
SDReturn_1	0.264	(0.12)	0.279	(0.13)
Price_1	-0.0247	(-0.37)	-0.0173	(-0.26)
Observations	301		301	

`T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Cox Regression for Bankruptcy Emergence with SD Rating Categories								
Main Effects								
careeropps_sd_2	**0.323***	**(1.97)**						
compensation_sd_2			0.144	(0.85)				
culture_sd_2					**0.491****	**(2.90)**		
seniorleadership_sd_2							**0.334***	**(2.00)**
Controls								
EBITTA_2	1.352	(1.67)	1.514	(1.87)	1.220	(1.50)	1.286	(1.57)
WCTA_2	-0.123	(-0.24)	-0.117	(-0.23)	-0.111	(-0.22)	-0.163	(-0.32)
TLTA_2	0.110	(0.56)	0.132	(0.68)	0.133	(0.68)	0.111	(0.57)
CHIN_2	0.0403*	(1.97)	0.0410	(1.96)	0.0381	(1.89)	0.0390	(1.89)
ExcessReturn_2	-0.0469	(-0.37)	-0.0704	(-0.56)	-0.0332	(-0.26)	-0.0690	(-0.56)
SDReturn_2	3.011	(1.52)	3.196	(1.65)	2.953	(1.48)	2.911	(1.46)
Price_2	0.0375	(0.54)	0.0445	(0.64)	0.0373	(0.54)	0.0445	(0.64)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* $p<0.05$
** $p<0.01$
*** $p<0.001$"

	Cox Regression for Bankruptcy Emergence with SD Rating Categories			
Main Effects				
worklife_sd_2	**0.395***	**(2.33)**		
overall_sd_2			0.286	(1.73)
Controls				
EBITTA_2	1.292	(1.59)	1.349	(1.66)
WCTA_2	-0.140	(-0.27)	-0.167	(-0.32)
TLTA_2	0.119	(0.61)	0.129	(0.67)
CHIN_2	0.0385	(1.88)	0.0397	(1.92)
ExcessReturn_2	-0.0321	(-0.26)	-0.0550	(-0.44)
SDReturn_2	2.919	(1.47)	3.032	(1.54)
Price_2	0.0387	(0.56)	0.0443	(0.64)
Observations	301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Cox Regression for Bankruptcy Emergence with SD Rating Categories								
Main Effects								
careeropps_sd_3	**0.598****	**(3.17)**						
compensation_sd_3			**0.477***	**(2.46)**				
culture_sd_3					**0.737*****	**(3.58)**		
seniorleadership_sd_3							**0.402***	**(2.07)**
Controls								
EBITTA_3	0.833	(1.50)	0.878	(1.59)	0.837	(1.51)	0.907	(1.65)
WCTA_3	-1.094	(-1.95)	-1.075	(-1.92)	-1.220*	(-2.10)	-0.985	(-1.75)
TLTA_3	-0.0438	(-0.17)	-0.0281	(-0.11)	-0.0854	(-0.33)	0.00563	(0.02)
CHIN_3	-0.00341	(-0.64)	-0.00335	(-0.64)	-0.00328	(-0.61)	-0.00321	(-0.61)
ExcessReturn_3	-0.282	(-1.95)	-0.284	(-1.95)	-0.262	(-1.80)	-0.284	(-1.95)
SDReturn_3	1.831	(0.62)	2.071	(0.72)	2.978	(1.02)	1.946	(0.68)
Price_3	0.146*	(2.13)	0.141*	(2.06)	0.154*	(2.25)	0.138*	(2.02)
Observations	301		301		301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

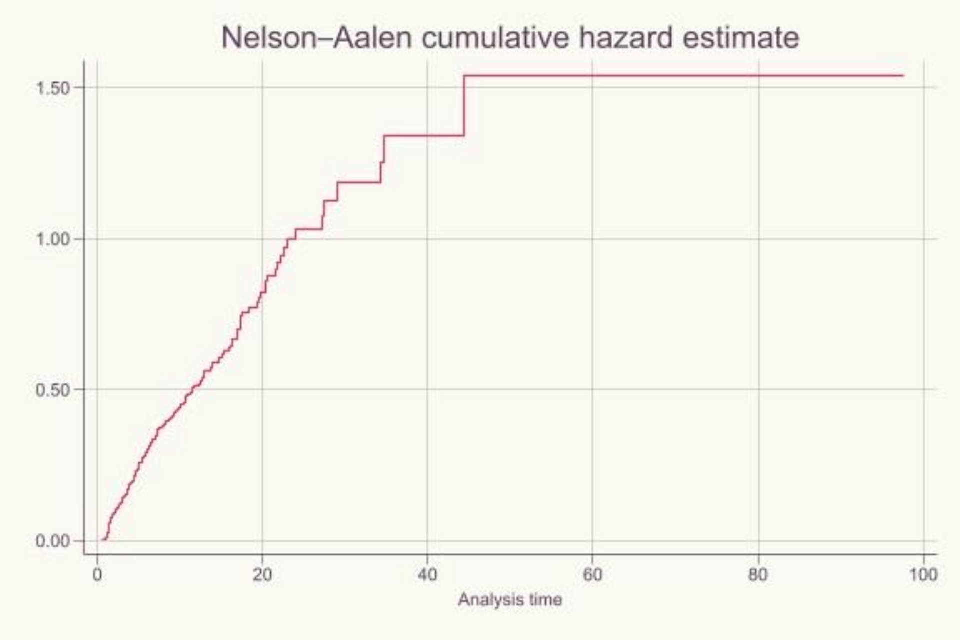
	Cox Regression for Bankruptcy Emergence with SD Ratings			
Main Effects				
worklife_sd_3	**0.465***	**(2.38)**		
overall_sd_3			**0.640*****	**(3.33)**
Controls				
EBITTA_3	0.886	(1.62)	0.856	(1.53)
WCTA_3	-1.068	(-1.88)	-1.178*	(-2.06)
TLTA_3	-0.0102	(-0.04)	-0.0675	(-0.26)
CHIN_3	-0.00331	(-0.63)	-0.00352	(-0.66)
ExcessReturn_3	-0.295*	(-2.02)	-0.290*	(-2.00)
SDReturn_3	2.058	(0.71)	1.866	(0.63)
Price_3	0.140*	(2.04)	0.151*	(2.19)
Observations	301		301	

T-Statistics in Parentheses
="* p<0.05
** p<0.01
*** p<0.001"

Figure A1: Cumulative Hazard Curve

Figure A1 presents Nelson-Aalen cumulative hazard estimate (Panel A) and Kaplan-Meier survival estimates (Panel B). Nelson-Aalen cumulative hazard estimate figure plots the analysis time (time from bankruptcy filing to time of emergence) on the horizontal axis, while the Kaplan-Meier survival estimates figure plots the analysis time (time from bankruptcy filing to time of emergence) on the horizontal axis.

<u>A1.1. Nelson-Aalen Cumulative Hazard Estimate</u>



__A1.2. Kaplan-Meier Survival Estimates__

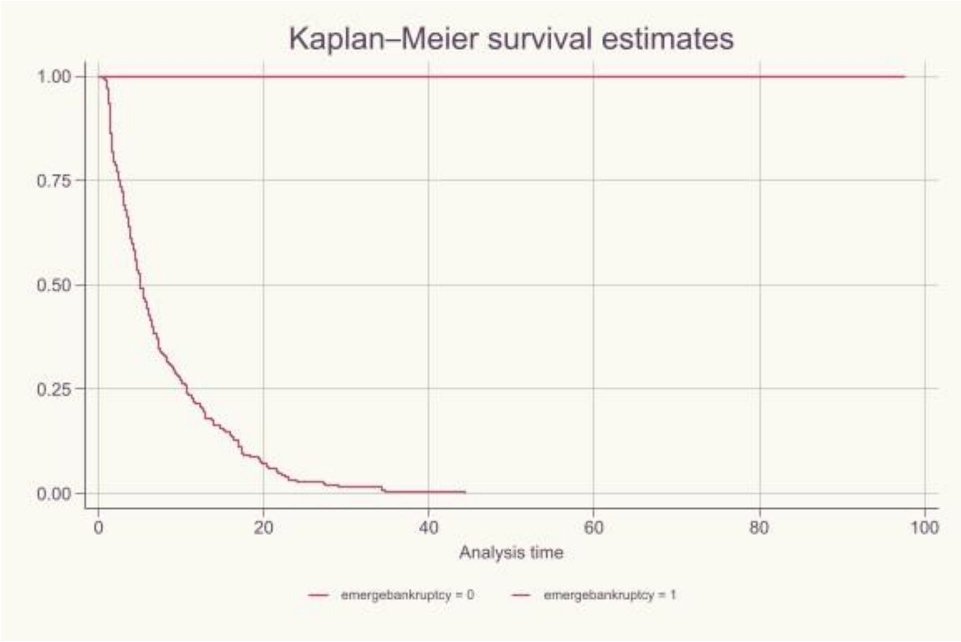
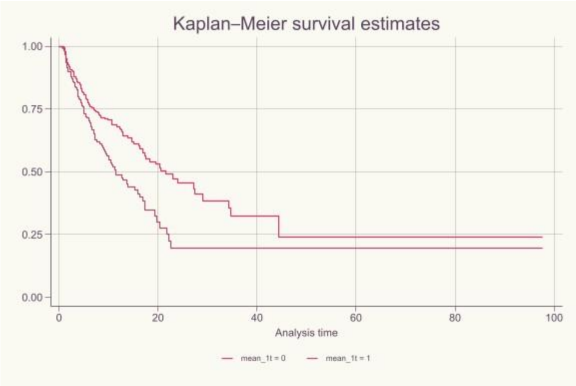
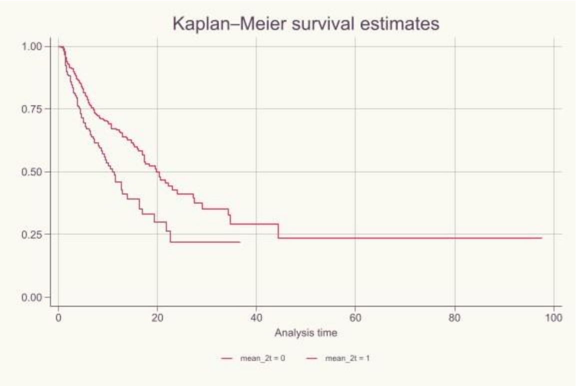


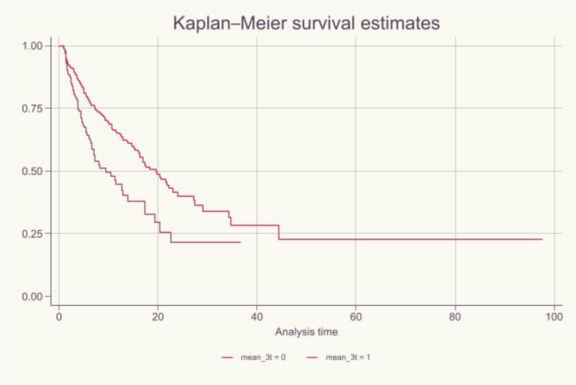
Figure A2: Univariate Analyses

Figure A2 presents Kaplan-Meier survival estimates from univariate analyses with mean employee satisfaction one, two, and three years before merger announcement (Panel A), standard deviation of employee satisfaction one, two, and three years before merger announcement (Panel B), and Z-Score one, two, and three years before merger announcement (Panel C).

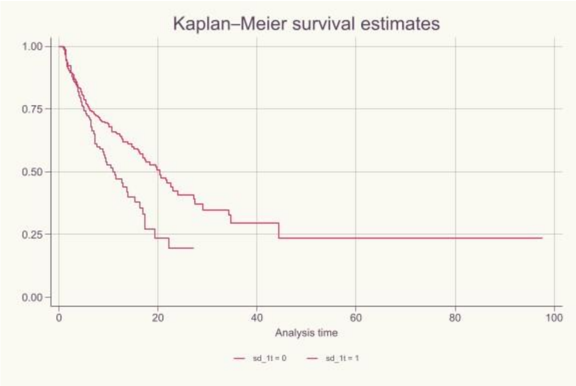
<u>A2.1. Kaplan-Meier Survival Estimates for Models with Mean of Employee Satisfaction</u>

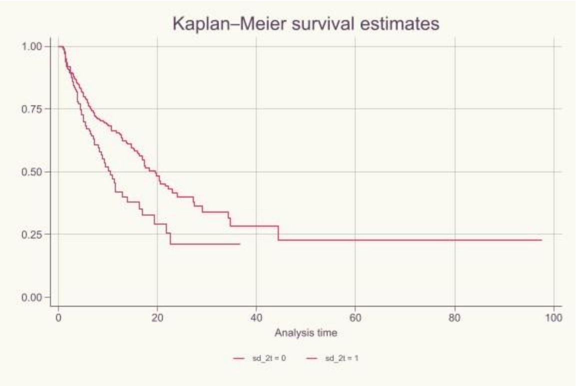


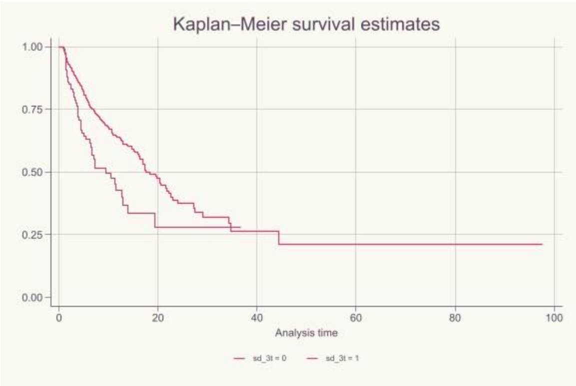




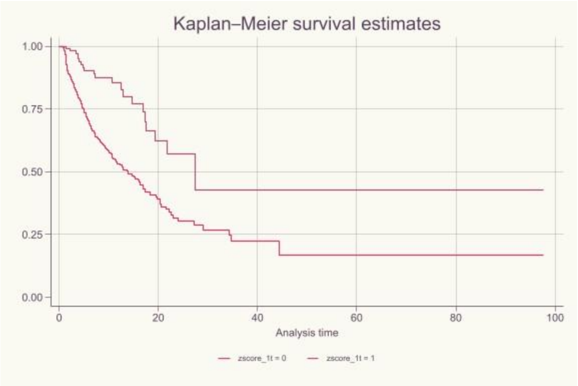
A2.2. Kaplan-Meier Survival Estimates for Models with Standard Deviation of Employee Satisfaction

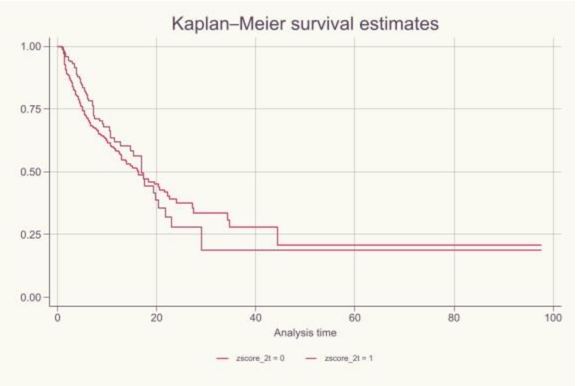


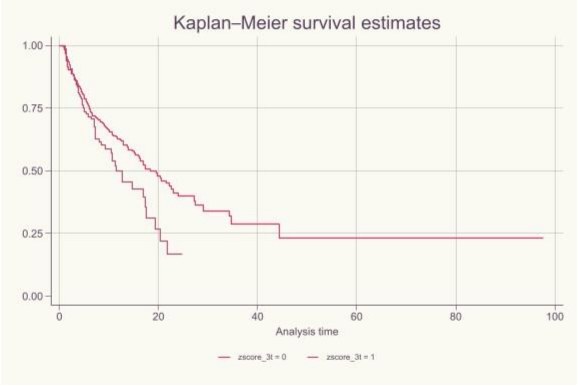




<u>A2.3. Kaplan-Meier Survival Estimates for Models with Z-Score</u>







Variable	Definition
EBITTA	(EBIT to Total Assets): Measures the firm's profitability relative to its assets. A negative value indicates that the firm is not generating enough profit to cover its operational costs.
Sales	Represents the total sales relative to total assets. Higher sales indicate greater efficiency in using assets to generate revenue.
NITA	(Net Income to Total Assets): Indicates profitability by comparing net income with total assets. Negative values indicate losses.
CHIN	(Change in Net Income): Measures the change in net income from the previous period, helping capture recent improvements or deteriorations in financial health.
WCTA	(Working Capital to Total Assets): Measures liquidity, indicating how well the firm can cover its short-term liabilities with its current assets.
CLCA	(Current Liabilities to Current Assets): Provides insight into liquidity, where higher ratios may indicate potential liquidity issues.
FUTL	(Funds from Operations to Total Liabilities): Reflects a firm's ability to meet its liabilities from operating cash flows.
INTWO	A binary variable that flags firms with negative income over the past two years.
RETA	(Retained Earnings to Total Assets): Measures how much profit the firm reinvests in its operations relative to its total assets.
MVETL	(Market Value of Equity to Total Liabilities): Represents the firm's leverage by comparing the market value of equity with total liabilities.
TLTA	(Total Liabilities to Total Assets): A common leverage ratio that indicates the proportion of a company's assets financed by liabilities.
OENEG	A binary variable indicating when a firm's liabilities exceed its assets, which could signal insolvency.
OhlsonSize	The logarithm of total assets adjusted for inflation, often used in Ohlson's bankruptcy prediction model.
RelativeSize	Logarithm of the firm's outstanding shares multiplied by share price, reflecting firm size in market terms.
Price	The log of the closing price at the end of the previous

	fiscal year, used as a market-based variable.
ExcessReturn	The firm's stock return in the previous year minus the market return (CRSP Index), used to gauge market sentiment.
SDReturn	The standard deviation of the residuals produced by regressing firm return on market return over the time period (usually 12 months). This value represents the SDReturn and captures the firm-specific risk or volatility in returns that is not explained by the overall market.
FirmAge	Age of each company relative to the time of incorporation.
Mean	The average of the sum of aggregated rating subcategories, or the average of individual rating subcategories one, two, and three years before a bankruptcy filing, which is used as a proxy for a company’s employee satisfaction average.
SD	The variation or dispersion of aggregated rating subcategories, or the variation or dispersion of individual rating subcategories one, two, and three years before a bankruptcy filing, which is used as a proxy for a company’s employee satisfaction variability.
careeropps_mean	Average of a company’s career opportunities ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
careerops_sd	Variability of a company’s career opportunities ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
compensation_mean	Average of a company’s compensation benefits ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
compensation_sd	Variability of a company’s compensation benefits ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
culture_mean	Average of a company’s culture values ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).

culture_sd	Variability of a company's culture values ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
seniorleadership_mean	Average of a company's senior leadership ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
seniorleadership_sd	Variability of a company's senior leadership ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
worklife_mean	Average of a company's work-life balance ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
worklife_sd	Variability of a company's work-life balance ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
overall_mean	Average of a company's overall rating ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
overall_sd	Variability of a company's overall rating ratings as provided in Glassdoor in the one, two, and three years before a bankruptcy filing (based on the test conducted).
Z-Score	A bankruptcy prediction score that uses a weighted sum of five financial ratios (working capital/total assets, retained earnings/total assets, EBIT/total assets, market value of equity/book value of liabilities, and sales/total assets) to assess a firm's financial health, where a lower score indicates a higher risk of bankruptcy.