ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Master Thesis Financial Economics
Title thesis: The predictive power of the Altman, financial and operational variables on financial distress for health care institutions
Name student: Jorrèl Abdoelhafiezkhan
Student ID number: 423671

Supervisor: Professor L. Swinkels
Second assessor: Professor R. Huisman
Internship supervisor: P. Diepstraten
Date final version: 30-07-2020

Abstract
This study tries to predict the probability of financial distress with the help of the Altman model and the logistic regression. A lot of studies concern bankruptcy, but financial distress (FD) has not received as much attention. The variables used for this study are divided in three groups namely the Altman variables originative from the Altman papers and the financial and operational variables coming from Finance Ideas. The models had a lower prediction accuracy than earlier research. The highest rate for correctly classified institutions is 75% for the operational variables in the full sample and 77.61% for the combined variables in the second period of the sub sample.

Index Page

# 1. Introduction

On the 25th of October 2018, the Slotervaart and Ijsselmeer hospital went bankrupt. This led to a lot of internal and external panic. The biggest problem was that patients whom were in critical conditions had to be transported to another hospital nearby. The people who lived around the hospitals were also in shock. They had to sign up at other hospitals to get their future treatments.

When the health care institutions (HCI) were taken over by their current owners, they were on the brink of bankruptcy, and could only be saved by a potential takeover. Although everyone knew that the hospitals were not financially healthy, they were naïve in handling the situation. It is important to learn from this incident so that similar problems will not occur in the future.

This research is aiming to predict which of the Altman, financial and operational variables increase or decrease the chance of financial distress (FD) for HCI. Previous studies attempted to predict bankruptcy using data of firms that went bankrupt. However, in this study, the amount of bankrupt HCI is insufficient, which is why financial distress is used as a proxy for bankruptcy.

For stakeholders, such as banks, it is important to know the financial status of an institution before providing the institution a loan, so the probability that the bank will lose money on this deal will decrease. If the financial department is informed about which characteristics influence the probability of financial distress, they can intervene on time to prevent a potential downfall of the HCI. In this thesis, I will contribute to this by answering the following research question:

*Is it possible to predict the deterioration of the financial status for health care institutions, that could potentially lead to financial distress, by using methods like the Altman's multiple discriminant analyses and the logistic regression?*

To support the research question, the following two sub-questions will be answered:

*Which variables have the most impact on financial distress?*

*Do the Altman, financial and operational variables differ over the years for healthy and non-healthy institutions?*

The five most used predicting methods are: (1) multiple discriminant analysis, (2) logit models, (3) neural networks, (4) contingent claims, and (5) univariate analysis (Altman et al., 2017). In this research the two most popular techniques will be used to predict financial distress.

Altman (1968) and Ohlson (1980) were two of the more well-known researchers that used models to predict the chance of bankruptcy. Their models were later used by a lot of institutions that needed to predict risk e.g. bankers, investors, asset managers etc. Besides using these methods in the banking environment, the methods could also be pleasing for investors and asset managers in helping them add healthy health care institutions to their portfolio.

Predicting bankruptcies is something that has been going on for quite some time. The classical papers about bankrupt firms in general are older than 50 years. The authors use techniques like the multivariate regression to predict the financial status of the firm. In this paper, the data of HCI is obtained through Finance Ideas. The data consists of a lot of variables of which I will use 17, plus the Financial Distress lag. The variables chosen are based on previous research and those that Finance Ideas use for their clients. The methods that apply to this data are the Altman Z-score model and the logistic regression. The results are that the Altman model was very good in predicting HCI in FD, whereas the logistic regression was better in predicting HCI not in FD. The rate for correctly classified institutions for the combined logistic regression in the second period (2014-2018) was 77.61%. This prediction rate is lower than previous research, but still has a good prediction accuracy. It's better to use this model than to take a guess, where the odds of one being correct is about 50%.

The remainder of the thesis will be as follows. In Section II, earlier research about bankruptcy, and financial distress in general will be discussed. After that I will look at more recent articles with the emphasis on bankruptcies on health care institutions. Following the literature, the data of previous research and Finance Ideas will be reviewed in Section III. Section IV describes the methodology used for testing the hypotheses. Section V will be about the results and Section VI will conclude with some ideas for future research & limitations.

# 2. Literature

It is important to discuss the underlying theory before starting with the research. The most well-known articles that describe the prediction of bankruptcy are Beaver (1966,1968), Altman (1968) and Blum (1969). Table 1 in the appendix gives a clear overview of the previous research that has been done on this topic.

Edmister (1972) states that Beaver's study is the basis for future research on ratio analyses. Beaver (1966) used a univariate discriminant analysis and focused on big asset-sized firms which failed between 1954-1964. He also collected firms that are comparable to the bankrupt firms. The most important variable Beaver found was annual cashflow over total debt. With this variable, he misclassified only 13 percent, one year before the actual bankruptcy. He mentions that there was approximately an equal number of failed and non-failed firms in the dataset, which means that the random prediction should be around 50 percent. He reasons that there is an extremely small probability, that a random prediction could perform as good as the analyses he used. With this finding, Beaver had high hopes for financial ratios in predicting business failure for at least five years before the event took place. Beaver (1968) concludes that liquid asset measures do a better job at predicting failures than nonliquid asset measures, even in the years that come immediately before failure. Liquid asset measures are financial indicators that are a proxy for liquidity, for example, the current ratio or the quick ratio.

Blum (1969) defined failure as "entrance into a bankruptcy proceeding or an explicit agreement with creditors which reduced the debts of the company". He collected 115 industrial firms based on financial and accounting data. The companies used in the sample failed in the period 1954-1968. Furthermore, he collected 115 comparable non-failing firms similar in industry, annual sales, number of employees, and fiscal year. The predictive accuracy of Blum's model for one, two and five years prior to bankruptcy are respectively 93-95%, 80% and 70%.

## 2.1. Discriminant analyses

#### 2.1.1.1. Altman's Z-score model

Since 1930, researchers applied the Multiple Discriminant Analyses (MDA) on their ratios. In the past, MDA was frequently used in the biological and behavioral sciences. This technique is primarily used to place an observation into a group that is defined beforehand, dependent on the individual characteristics of an institution. It is mainly used in situations where the dependent variable has a qualitative form, like FD in this study. An HCI can be in financial distress or not. “MDA then attempts to derive a linear combination of these characteristics which ‘best’ discriminates between the groups.” Altman (1968).

In 1968, Altman used financial ratios to predict the probability of a firm going bankrupt. He states that financial ratios are very important in predicting the institution’s performance. Before there were quantitative measures of the performance of a company, agents needed to give qualitative information to determine the creditworthiness of a particular merchant (Foulke, 1961). Altman explains that in previous research, firms that were in financial distress displayed different ratio measurements than firms that continued their operational activities (Merwin, 1942). According to Altman (1968), ratios measuring profitability, liquidity and solvency proved to be the best indicators. He used the following ratios for his research: working capital over total assets (WCTA), retained earnings over total assets (RETA) earnings before interest and taxes over total assets (EBITTA), market value of equity over book value of total debt (MVTL), and sales over total assets (STA). In total he tested 66 firms between 0.7 and 25.9 million dollars, in the period from 1954-1968. Of these 66 firms, 33 firms went bankrupt and the other 33 were comparable healthy firms. Altman (1968) used an F-test to review the correlation between the variables. The predictability of Altman’s model was 79 percent one year before the bankruptcy. Deakin (1972) improved Altman’s model by using 14 financial ratios, in the period 1964-1970. The error rates he found, were respectively 22%, 6%, 12%, 23% and 15% for one-five years before the event took place.

Previous studies used univariate regression models, but according to Altman these are questionable. Ratio analyses presented in this way is affected by wrong interpretation. “A firm with a poor profitability and/or solvency record, may be regarded as potentially bankrupt. However, because of its above average liquidity, the situation may not be considered serious.” (Altman, 1968). Altman used a multiple discriminant analyses (MDA), also called the Altman Z-score model, which allowed him to combine multiple ratios and have a more complete answer than a standard univariate regression. Even though there are other models that have a better predicting power than the Altman’s Z-score model, Altman still uses this model because it is accounting-based. This means that it only needs accounting data for the prediction.

According to Agarwal & Taffler (2008), the accounting-ratio-based approach has three parts in its favor. These accounting techniques give a good overview of a company. A healthy company with a good financial statement will not suddenly go bankrupt, because of a change in the economic environment. Most of the time, the accounting statements of firms will capture the financials, so corporate failure is usually the end result of several years of negative performance. Because of the double entry system, changes in the accounting policies or window dressing the accounts will have minimal effects on the financial statement. "Finally, loan covenants are generally based on accounting numbers and this information is more likely to be reflected in accounting-ratio-based models." (Agarwal & Taffler, 2008).

Altman et al. (1976) collected 40 bankrupt firms and 113 healthy firms between 1970-1972. He also used the MDA in this study and found that the prediction accuracy for failed, and healthy firms were 90 and 90.3 percent. The type I error is almost the same as the type II error and is respectively 10 and 9.7 percent. This is a good prediction accuracy, because the model can correctly classify an HCI in 90% of the cases.

Gu and Gao (2000) used 28 firms, divided in two groups, to predict the bankruptcy. Each group contained 14 firms and these groups were split in bankrupt and non-bankrupt firms. The prediction period was from 1987 to 1996 and the firms were grouped based on total assets and standard industrial classification (SIC) code. Gu and Gao made their predictions with the help of the MDA model and five variables. The variables they used were, total liabilities over total assets, EBIT over current liabilities, gross profit margin, long-term debt over total assets, and sales over fixed assets. The predictive accuracy of the model was 93% one year beforehand.

Other articles that use the Altman model to predict bankruptcy are Russ et al. (2004), and Dambolena and Khoury (1980). The Type I and Type II error of Russ et al. (2004) are respectively 20.6% and 28.4%. The prediction accuracy of Dambolena and Khoury were 87%, 85% and 78% for one, three and five years prior to the bankruptcy. They used 68 firms in their data set between 1969-1975.

### 2.1.2. *Revised Z-model for private firms*

The original Altman's Z-score model was developed for public companies. Altman et al. (1983) adjusted his Z-model, so it was suitable for private firms. The coefficients, variables and the identification zones were adjusted for this new model. The sample requirements are: firms need to be unlisted industrial (non-financial) companies, limited owner liability, big enough company size and finally, the countries need at least 60 companies in financial distress, or they will only be added to the test sample. Altman was able to correctly predict 90.1% of bankrupt firms and 97% of healthy firms in his 1983 paper.

In the research of Altman et al. (2014), 34 countries were used in the data sample, of which 31 were European and three were non-European countries. Furthermore, the firms in the sample were private and non-financial companies across all industries. Because of this, the revised Altman model of 1983 was used. According to Altman et al. (2014) the predictive power was not equal for the countries. In some countries the model did exceptionally well, whereas in others it underperformed. For most of the countries, the predictive power of the revised Altman model was around 75%. Some countries did exceptionally well with prediction accuracy levels above 90%.

According to Altman (1993), the revised model varies in prediction accuracy over different industries. In the first 20 years the main focus was on manufacturing firms that differed in size. Although it is one of the most well-known models for predicting bankruptcies, Altman advises to combine this model with other models, and not solely rely on the Altman model.

In the paper of Altman et al. (2014), the authors conclude that the revised Altman model is one of the best Multiple Discriminant Analyses one can use. The model has been researched for over 40 years and is used in different industries/sectors to predict the bankruptcy. The industries/sectors that researchers analyzed were, the service and manufacturing industry, publically listed companies (only for the original Altman model), and banks. All Altman's models have been used in different researches with reliable outcomes, therefor it can be safely said that for one, two and three years prior to FD or bankruptcy, the three Altman models can be used.

**Table 2: Previous Altman studies**

The table shows a nice overview of the different Altman models that were developed over the years. For the original model $X_1$ to $X_5$ are respectively Working capital/ Total assets, Retained earnings/Total assets, Earnings before interest and taxes/ Total assets, Market value of equity/Book value of total liabilities, and Sales/Total assets. In the revised model, market value of equity is replaced with book value of equity and finally for the revised four model the Sales/Total assets variable is removed from the formula. Anjum (2012).

| *Coefficient Variables* | *Original Model (1968)* | *Revised Model (1983)* | *Revised Four Model (1993)* |
|---|---|---|---|
| $X_1$ | 1.21 | 0.717 | 6.56 |
| $X_2$ | 1.41 | 0.847 | 3.26 |
| $X_3$ | 3.30 | 3.107 | 6.62 |
| $X_4$ | 0.60 | 0.42 | 1.05 |
| $X_5$ | 0.999 | 0.998 | N/A |
| *Cut-off scores* | | | |
| *Non-Bankrupt Firms* | <1.81 | <1.23 | <1.10 |
| *Gray Area* | 1.81-2.67 | 1.23-2.90 | 1.10-2.60 |
| *Bankrupt Firms* | >2.67 | >2.90 | >2.60 |

### 2.1.3. *Critique on the MDA*

As mentioned in the introduction, the MDA and the logistic regression are the most popular techniques for predicting bankruptcies. Like any other model, these models also have their flaws. Ohlson (1980) mentions three reasons why his logistic regression is better than the MDA model. First, for the distributional properties of the independent variables there needs to be certain statistical requirements. One of these requirements is that the variance-covariance matrix of both groups (bankrupt and non-bankrupt firms) has to be the same. According to Ohlson (1980) it could be argued that violation of this requirement is not important if the model is only used for discriminating purposes. Second, it is hard to interpret the output of the MDA model since it is a score of an ordinal ranking system. “For decision problems such that a misclassification structure is an inadequate description of the payoff partition, the score is not directly relevant.” Ohlson (1980). Finally, MDA falls victim to some “matching” procedures.

Criteria like size and industry are considered random when they are matched with failed and non-failed firms. Anyhow, it seems to be better if variables are included as predictors instead of using them for matching purposes.

At last, Altman (1983) considered the implementation of his Z-score model questionable. "He admitted that the model did not scrutinize very large and very small firms, the observation period was quite long (almost two decades), and the analysis included mostly manufacturing companies." (Altman, 2014). In conclusion Altman wanted a model that could predict bankruptcy with a homogenous group of bankrupt firms and the newest available data.

## 2.2. Logistic regression

When applying the conditional logit analysis, all problems that arise with using the MDA can be avoided. "No assumptions have to be made regarding prior probabilities of bankruptcy and /or the distribution of predictors." (Ohlson,1980).

Ohlson (1980) was one of the first scientists that used logistic regressions to predict bankruptcy. He found that the size of the company, financial structure, performance, and liquidity had a statistically significant effect on the probability of bankruptcy. He used a data sample of 105 bankrupt firms and 2058 non-bankrupt firms in the period from 1970 to 1976. The predictive accuracy of his model was 92% two years earlier. Previous research used financial statements after the firm went bankrupt to create their model. This gives a clear upward bias for the accuracy. Even though this factor is implemented in this research, the predictive error is still larger than the study of Altman's (1968) and Altman et al. (1977). The results of Altman and McGough (1974) and Moyer (1977) are more in line with the results of this paper.

Ohlson (1980) used size (log(total assets/GNP price-level index), total liabilities divided by total assets, working capital divided by total assets, current liabilities divided by current assets, a dummy variable were the value is one if total liabilities exceeds total assets and zero otherwise, net income divided by total assets, funds provided by operations divided by total liabilities, a dummy were the value is one if the net income was negative the last two years, zero otherwise, and a change in net income to predict bankruptcy.

Zavgren (1988) also used the logistic regression to calculate the prediction accuracy of the variables one to five years before the event. The total classification errors for 1 to 5 years prior to bankruptcy were 18%, 17%, 28%, 27% and 20%. The data sample was obtained in the period from 1979-1980. In the short run his results were similar to that of Ohlson his logistic regression. In the long term his predictive accuracy was similar-to-slightly lower than studies that applied the MDA. Darayseh et al. (2003) studied 220 firms in total from 1990-1997. He split the 220 firms in 110 bankrupt and 110 non-bankrupt firms matched by total assets and industry. The prediction power of the model for the in-sample and the holdout sample were respectively, 87.82% and 89.50%.

Lau (1987) analyzed five states to test the probability of bankruptcy. The five states all had different criteria for testing. Each sample contained 350 firms. The sample period was between 1971 and 1977. Lau found that the original sample had a prediction accuracy of 96%, 92% and 90%, one, two and three years before the event took place. For the holdout sample the probabilities were 80%, 79% and 85% for one, two and three years respectively. She did not mention anything about a non-clear trend in the prediction accuracies for the holdout sample.

Kim & Gu (2006) developed a logit model to predict the bankruptcy in the hospitality industry. Between 1999 and 2004 they collected 32 firms in total, 16 bankrupt firms and 16 matching non-bankrupt firms and tried to predict the bankruptcy up to two years in advance. Their logit model could correctly predict 91% and 84% of the bankruptcies for one and two years prior to the event. The predictive power is similar to earlier studies. Their key finding is the cashflow to total liabilities just like Beaver in 1966. A firm that highly relies on debt financing and has trouble generating cashflow is more vulnerable for bankruptcy.

### *2.2.1. Critique on the logistic regression*

The logistic regression is a flexible method to determine the probability of bankruptcy. It does not contain a lot of requirements and restrictions. "Logit regression analyses does not assume multivariate normality, but is, however, exposed to a full linear compensation between variables in the exponent of the logistic function." Laitinen and Laitinen (2000). Collins and Green (1982) state that the logit model is not much better in forecasting as one might expect beforehand. The other forecasting models like described in Altman (1968,1973) and Altman (1977) have almost similar results.

## 2.3. Bankruptcy for hospitals

There are only a few studies about hospital/ health care institutions in financial distress. Morey et al. (2004) and Almwajeh (2004) focused on predicting financial distress and bankruptcy for hospitals.

In their research, Morey et al. investigated five upstate New York hospitals. They wanted to determine a frequent pattern that might predict future financial distress, bankruptcy or even closure. Morey et al. (2004) did this by taking four years of data from each of the hospitals that went either in distress, bankrupt or closed down. The financial variables used for his study were liquidity (Cash Flow to Total Debt), leverage Long-term debt per bed, profitability (Equity Financing) and efficiency. They found that the main reasons for discontinuation of an entity were: business related decisions (relocation, consolidation, or mergers), low number of patients, increase of costs and/or decrease of revenue, and finally increasing competition.

Almwajeh (2004) also tried to predict financial distress and bankruptcy using the financial ratios of 65 rural hospitals. The Altman's Z-score model with modified financial ratios of liquidity, profitability, efficiency and productivity was applied for this research. "Liquidity and profitability ratios have the highest contribution to the results of the Z-scores, followed by productivity and efficiency." (Almwajeh, 2004). Almwajeh's concluded that both the Discriminate Analyses and Logistic Regression models are able to predict the financial status of the hospitals with a prediction power of respectively 90.2% and 100%.

# 3. Data

The required data is obtained through Finance Ideas (FI). FI has a database called "ZorgRating" in which they collected data of Health Care institutions between 2009-2018. The dataset is very broad in the number of HCI and the number of observations per HCI. There are 17 variables collected from the dataset which are collected for different companies over time. In the database there are 852 firms in total. Of these 852 firms, 818 were non-bankrupt firms and 34 were firms that went bankrupt. In total there are 7535 panel data observations, and the mean of the total assets is $66.55 million. In the dataset there is a column title called notes[1]. The notes column states the current situation of the HCI. An HCI can be marked as active[2] or inactive[3]. The notes show if an HCI is stil

[1] Notes = Notities in the dataset. [2] Active = Actief. [3] Inactive= Inactief.

functioning (Active[2]), if it went bankrupt or is acquired by, or merged with another HCI (Inactive[3]). The institutions that did not have fixed assets[4], current assets[5], equity[6], provision[7] and long-term debt[8] were removed from the sample.

For this research, the focus is on FD as a proxy for bankruptcy. A dummy for FD was created were one indicates an HCI in FD, and zero an HCI not in FD. Approximately 0.44% of the institutions in the dataset went bankrupt and a little more than 45% of the health care institutions were in financial distress. The 45% is a high percentage, this indicates that 45% of the firms throughout the 10 years of observations, had a DSCR or solvency lower than the minimum set by the banks. It is noteworthy that such a big amount of the institutions is in financial distress and only 0.44% of these firms went bankrupt over the years. The reason for this difference is that FD is noted per year. For example, if institution A was in FD in 2009, 2011 and 2015, this would be noted 3 times as FD, but institution B can only go bankrupt one time. An institution that went bankrupt usually was in FD the year before, but not the other way around. There are also institutions that go into FD and never go bankrupt. FD is not the perfect proxy to predict bankruptcy, but it is the beginning stage. If we can prevent HCI from going into FD, or intervene as soon as possible, the probability of future bankruptcies could be reduced.

Altman used 5 variables for his Z-score model, and Finance Ideas uses 16 variables to predict the financial position of health care institution. Of these 16 variables, 8 are financial variables and 8 operational variables. Two financial ratios, solvability ratio and Debt Service Coverage Ratio (DSCR), and two operational ratios (customer satisfaction and non-salaried employees) are left out of the thesis. The solvability ratio and DSCR are left out of this study, because these two variables are used to classify a health care institution in financial distress or not. The customer satisfaction variable is left out because it is not consistently collected over time and non-salaried employment is only available for more recent periods and not for the complete dataset. The variables are divided into the following three categories: Altman variables for the Altman test and the logistic regression, and financial and operational variables for the logistic regression.

Banks take a look at financial covenants like the Debt Service Coverage Ratio (DSCR) and the Solvability of a company to decide if the institution is in FD or not. If the DSCR is below 1,4

[2] Active = Actief. [3] Inactive= Inactief. [4] Fixed Assets = Vaste Activa. [5] Current Assets = Vlottende activa. [6] Equity = Eigen Vermogen [7] Provisions = Voorzieningen. [8] Long-term debt = Lange-termijn schulden.

and/or the solvability is below 20% the institution is characterized as financially distressed. The calculations for the solvability and the DSCR are noted in the following formulas:

$$\text{Solvency rate} = \frac{After-tax\ net\ operating\ income}{Total\ debt\ obligations}$$

$$\text{DSCR} = \frac{EBITDA}{Interest\ income\ - the\ interest\ costs + repayment\ of\ the\ financial\ year}$$

All the variables that are used for this research are *winsorized*. Winsorization means that extreme values of a dataset are limited which reduces the effect of outliers. In other words, the smallest 1% and biggest 1% of the variable are capped.

Looking at the number of health care institutions in figure A there is an observable difference between 2018 and the years before. The main reason for the decrease in the number of institutions is because of mergers and acquisitions. An institution that changes its name is registered as a new institution. This data is reliable because it has been certified by an accountant. FI uses this data to make reports for their clients. Figure B shows a slight downtrend of firms in Financial Distress. A possible reason for this situation could be that health care institutions are more concerned about their financial position and take more precautions to keep FD from happening. In Figure C one can observe that the number of health care institutions that went bankrupt increased significantly in 2013.

Figure A.

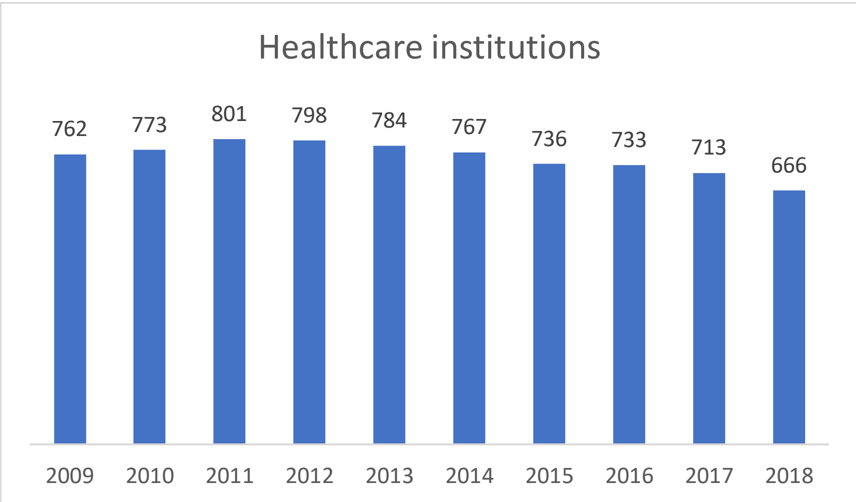

Figure B.

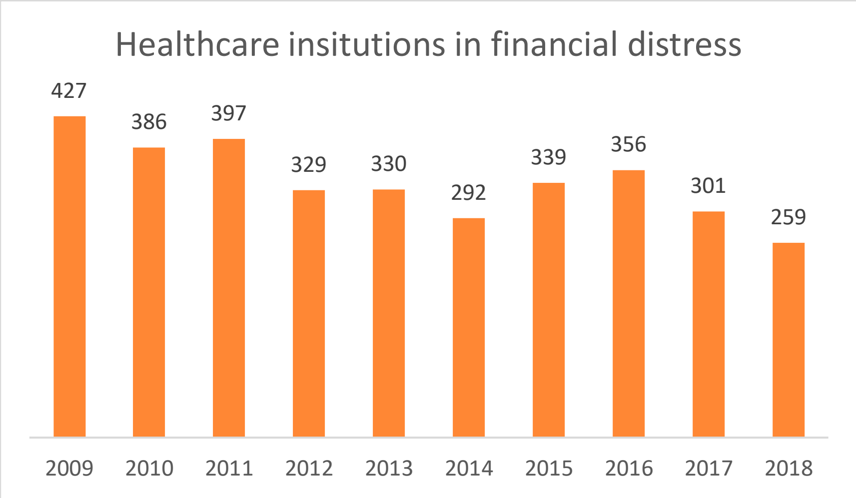


Figure C.

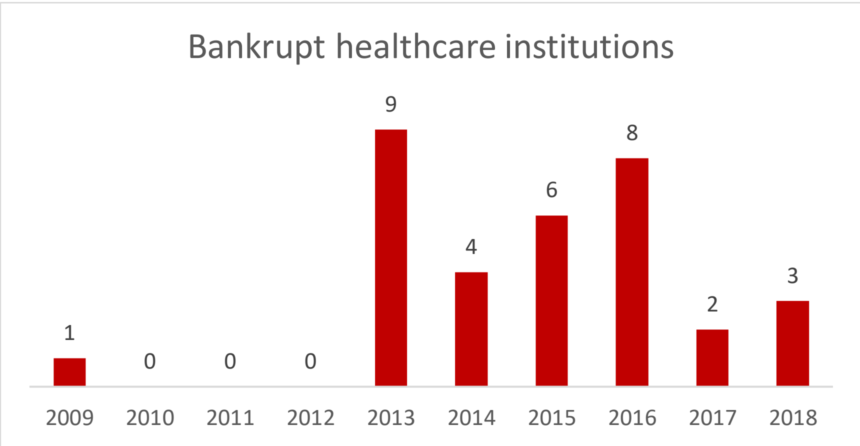


Table 3: Variables for the Altman and Logistic model

| *Revised Altman Z-score model & Logistic regression* | *Logistic regression* | |
|---|---|---|
| *Altman variables* | Financial variables | Operational variables |
| *Working Capital/ Total Assets* | Current ratio | Absenteeism |
| *Operating Results/Total Assets* | Efficiency ratio | Staff turnover |
| *EBIT/ Total Assets* | Result ratio | Competition |
| *Book Value Equity/ Total Liabilities* | Net debt/ EBITDA | Book value/ Purchase value |
| *Revenue/ Total Assets* | LnRevenue | Accounts receivable |
| | Revenue development | Annual statement |
| | Financial Distress lag | Financial Distress lag |

The retained earnings over total assets and sales over total assets were not available for the HCI. HCI don’t have retained earnings and sales, therefor, proxies of these variables are applied in the thesis. Instead of the retained earnings over total asset and sales over total assets, the operating results over total assets and revenue over total assets are used. These variables were the most identical to the original variables. The modification of the variables can lead to a better/worse prediction accuracy.

Variables for the revised Altman model (Altman,1983):

1. Working capital/total assets (WCTA)= Working capital is the difference between current assets and current liabilities. This ratio is a good liquidity measure. A firm that experiences losses, has lower current assets compared to fixed assets.
2. Operating result/total assets (ORTA) = Operating result is the difference between operating revenue and operating costs. A healthy company has their finances under control and a higher operating result compared to FD companies.
3. EBIT/Total assets (EBITTA) = This is a good productivity measure. The ratio indicates the earning power of the assets in place.
4. Book value equity/Book value of total liabilities (BVETL)= The original model has market equity/book value of total liabilities as the fourth variable. Private companies don't have market equity, that is why book value of the equity is a good replacement.
5. Revenue/Total assets (RTA) = Hospitals don't have a post for sales. Revenue is a good replacement, because the number of sales times the sell price of a product will eventually be the same as the revenue.

Financial variables for the logistic regression (Finance Ideas):

1. Current ratio (CR) = The ratio between the current assets and the current liabilities. The ratio indicates how well the institution is able to pay its short-term debt. A current ratio above 1 means that the company is liquid enough to pay its debt, and vice versa.
2. Efficiency ratio (ER) = The ratio between the operating costs and the operating revenue of the institution. The lower the ratio, the better the institution operates, because the institution needs less expenditures for a certain amount of revenue.
3. Result ratio (RR) = The ratio between the operating result and the operating revenue. A higher RR means that a firm is better able to minimize costs and maximize revenue.
4. Net debt/EBITDA (DEBITDA)= The ratio indicates the number of years a company needs to pay back the outstanding debt. If the ratio is high, the institution has trouble

paying back the outstanding debt. $\frac{net\ debt\ (long-term\ debt+debt\ to\ credit\ institutions-liquid\ assets)}{EBITDA\ (operating\ result+impairments+depreciation+changes\ in\ provision)}$

5. LnRevenue = The size of the revenue indicates the size of the company. In general, companies with a higher revenue can deal better with setbacks than companies that have lower revenues. It's easier for them to take in the lost revenue, because they have more. The natural logarithm is used because Revenue is not normally distributed. The Revenue's vary a lot from each another and the natural logarithm can help smooth the distribution.
6. Revenue development (RevenueD) = Is the development of the revenue over the previous years. If a company shows a steady growth of the revenue, it will be rated higher than a company with a low or negative revenue growth. Revenue D is calculated by taking the difference between t and t-1 and dividing the result by t-1. Finally multiply the outcome by 100%.

   RevenueD = $((Revenue_t - Revenue_{t-1})/Revenue_{t-1})$ *100%

Operational variables for the logistic regression (Finance Ideas):

1. Absenteeism = The number of employees that do not show up to work, because of illness. A higher rate means that there are more sick employees. This will make it more difficult for the institution to pick up the work that was intended for those employees.
2. Staff turnover (ST) = Staff turnover is the percentage of staff that's leaving the institution. A lower staff turnover indicates continuity, more stability and more certainty.
3. Competition = Shows the competition for an institution. The variable is calculated by dividing the number of patients by the number of institutions in the same postal code. A high number indicates enough patients for the area and a healthy/no competition.
4. Book value/Purchase value (BVPV) = Indicates if the institution needs a renovation for its buildings. If the book value is low compared to the purchase value, it's time for investing in renovation. This is a counterbalance variable for the solvency, after an investment, the solvency rate will be lower and this ratio will be higher. Firms that do not have their finances under control, usually have a lower BVPV.
5. Accounts receivable = This ratio shows if the firm claims their invoices on time. A firm that has a lot of outstanding invoices has less capital, and this can lead to higher risks on liquidity shortcomings.

6. Annual statement (AS)= This variable indicates if the institution is on time with disclosing their annual statement. An institution that discloses its annual statement earlier has more control in their organization. The annual statement variable is a dummy. The institution is late in disclosing its annual statement after the 1st of June. The dummy gets a value of 1 if the annual statement is disclosed after the 1st of June and 0 otherwise.
7. Financial distress lag = The lagged Financial distress variable tries to predict if the probability of FD increases, when an HCI was distressed in the previous period.

An institution needs to be active for at least three consecutive years to be part of the data. By taking this requirement into account it is possible to make more accurate predictions then if this is not the case.

The following tables have the same construction as Table 4 of the paper of Gu and Gao (2000). The means of the variables are divided into two categories. In the first category are healthy HCI and the second category FD HCI. After the means of both categories are calculated, the difference between the means are taken. This method is used to check if there is a significant difference between the two categories, (Deakin, 1972).

In Table 4, the variables for the revised Altman Z-score model are stated. All the variables are significantly different at the 1%-level between HCI in FD and those that are not in FD. The coefficients are in line with the expectations except for revenue over total assets. The revenue over total assets is higher for HCI in FD than for institutions that are not in FD. Revenue over total assets is a proxy, this can explain why the sign of the variable is not in line with the theory. The coefficients in Table 4 have the same sign as the paper of Gu and Gao (2000). The sign for working capital over total assets is the same as in Altman (1968), Gu and Gao (2000) and Ohlson (1980).

Table 4: Altman variables

| | *Not in Financial Distress* | | *Financial Distress* | | *Difference* |
|---|---|---|---|---|---|
| | Means | Std dev | Means | Std Dev | P-value |
| ***Altman variables*** | | | | | |
| *Working capital/Total Assets* | 0.202 | 0.244 | 0.041 | 0.250 | 0.000*** |
| *Operating results/Total Assets* | 0.069 | 0.086 | -0.005 | 0.089 | 0.000*** |
| *EBIT/Total Assets* | 0.199 | 0.090 | 0.120 | 0.111 | 0.000*** |
| *Book Value Equity/Total Liabilities* | 1.186 | 1.408 | 0.519 | 0.863 | 0.000*** |
| *Revenue/ Total Assets* | 1.562 | 0.913 | 1.752 | 1.305 | 0.000*** |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Table 5 shows a significant difference between HCI that were not in FD and those that were. All the financial variables showed a siginificant difference on the 1%-level except for LnRevenue and RevenueD. The outcomes are in line with the expectations. The operational variables are significantly different at the 1%-level except for book value over purchase value, this variable is significantly different at the 5%-level. The coefficients of competition and BVPV are not in line with the expectation. The competition variable shows that there are more patients for HCI in FD than for HCI that are not. This is contradicting because one would expect that institutions with less patients are more likely to be in FD. The coefficient for BVPV is higher for institutions in FD than for institutions not in FD. According to the theory, an institution that is in FD, has a lower BVPV, because the HCI does not have enough capital to renovate the buildings. If they have enough capital, renovating the building is likely one of the last items on their list. The difference between the current ratios in the paper of Gu and Gao (2000) is not significant on any level while the current ratio in this thesis is. The numbers for the Financial Distress variable are not available because the t-test is grouped by Financial Distress, it is the same variable. Most of the variables that are used in this paper were not studied in earlier researches.

Table 5: Financial and Operational variables for the logistic regression

| | *Not in Financial Distress* | | *Financial Distress* | | *Difference* |
|---|---|---|---|---|---|
| **Logistic regression variables** | Means | Std dev | Means | Std Dev | P-value |
| ***Financial variables*** | | | | | |
| *Current ratio* | 1.974 | 1.412 | 1.294 | 0.966 | 0.000*** |
| *Efficiency ratio* | 0.745 | 0.080 | 0.770 | 0.088 | 0.000*** |
| *Result ratio* | 0.044 | 0.045 | 0.001 | 0.425 | 0.000*** |
| *Net Debt/EBITDA* | 4.079 | 3.804 | 7.26 | 15.063 | 0.000*** |
| *Ln Revenue* | 17.124 | 1.511 | 17.012 | 1.529 | 0.016** |
| RevenueD (%) | 0.095 | 1.345 | 0.080 | 0.275 | 0.645 |
| ***Operational variables*** | | | | | |
| *Absenteeism* | 0.051 | 0.017 | 0.055 | 0.018 | 0.000*** |
| *Staff turnover* | 0.133 | 0.128 | 0.147 | 0.146 | 0.000*** |
| *Competition* | 8727.329 | 22342.77 | 11,868.53 | 26409.24 | 0.000*** |
| *Book value/Purchase Value* | 0.504 | 0.178 | 0.513 | 0.196 | 0.035** |
| *Accounts Receivable* | 0.028 | 0.042 | 0.035 | 0.050 | 0.000*** |
| *Annual Statement* | 0.288 | 0.453 | 0.358 | 0.480 | 0.000*** |
| *Financial Distress lag* | N/A | N/A | N/A | N/A | N/A |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Hand (1981) and Huberty (1984) argue that more variables don't necessarily lead to a better discriminant function. Having less variables, creates a simpler and easier to read model. Also, the model could perform just as well if compared to more variables.

Altman (2000) mentioned in his research that there is a possibilty for a variable bias, because of the intensive searching. He states that there is no evidence that a set of variables, proved to be effective in the initial sample, will be effective for the whole population.

## 3.1. Lagged variables

In this paragraph the development of the variables for one, two and three years before the event date will be discussed. The lags are for all observations, not only for the first time the HCI went into FD. In normal circumstances the difference between the variables need to increase over time. For example, the difference between the mean of the variables three years before the event date has to be smaller than the difference of a two year lag, and the difference between the mean of a two year lag needs to be smaller than a one year lag. Table 6 shows, that the development of the differences, are in line with the expectations, except for RTA. Here the difference between an HCI in FD and not in FD is decreasing over time. It is also noteable that all the differences are significant at the 1%-level.

In Table 7, The difference between HCI in FD and those that are not in FD increase, except for the ER, LnRevenue and the RevenueD. The difference in the ER is decreasing from lag 3 (0.011) to lag 2 (0.010) and then increasing from lag 2 (0.010) to lag 1 (0.013). The difference in the means of LnRevenue decreases over time. This is an odd event, because this indicates that revenue becomes less important, the closer the variable is to the event date. And finally, the difference in RevenueD is not significant for any lag.

Table 6: Lagged variables of the Altman model

| *Altman variables* | *Lag 1* | *Lag 2* | *Lag 3* |
|---|---|---|---|
| ***Working capital/ Total assets*** | | | |
| *Fin Distress* | 0.066 | 0.076 | 0.079 |
| *Non-Fin Distress* | 0.168 | 0.149 | 0.132 |
| *Difference* | -0.102*** | -0.073*** | -0.053*** |
| ***Operating results/Total assets*** | | | |
| *Fin Distress* | 0.011 | 0.022 | 0.027 |
| *Non-Fin Distress* | 0.057 | 0.052 | 0.053 |
| *Difference* | -0.046*** | -0.030*** | -0.024*** |
| ***EBIT/Total assets*** | | | |
| *Fin Distress* | 0.141 | 0.152 | 0.159 |
| *Non-Fin Distress* | 0.188 | 0.184 | 0.188 |
| *Difference* | -0.047*** | -0.032*** | -0.029*** |
| ***Book value/Total assets*** | | | |
| *Fin Distress* | 0.594 | 0.624 | 0.636 |
| *Non-Fin Distress* | 1.067 | 0.995 | 0.953 |
| *Difference* | -0.473*** | -0.371*** | -0.317*** |
| ***Revenue/Total assets*** | | | |
| *Fin Distress* | 1.701 | 1.690 | 1.698 |
| *Non-Fin Distress* | 1.573 | 1.558 | 1.549 |
| *Difference* | 0.128*** | 0.132*** | 0.149*** |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Table 7: Lagged variables of the financial variables

| *Financial variables* | *Lag 1* | *Lag 2* | *Lag 3* |
|---|---|---|---|
| ***Current ratio*** | | | |
| *Fin Distress* | 1.401 | 1.428 | 1.439 |
| *Non-Fin Distress* | 1.821 | 1.742 | 1.683 |
| *Difference* | -0.420*** | -0.314*** | -0.245*** |
| ***Efficiency ratio*** | | | |
| *Fin Distress* | 0.762 | 0.759 | 0.757 |
| *Non-Fin Distress* | 0.749 | 0.749 | 0.746 |
| *Difference* | 0.013*** | 0.010*** | 0.011*** |
| ***Result ratio*** | | | |
| *Fin Distress* | 0.010 | 0.015 | 0.018 |
| *Non-Fin Distress* | 0.038 | 0.035 | 0.036 |
| *Difference* | -0.028*** | -0.020*** | -0.018*** |
| ***Net debt/EBITDA*** | | | |
| *Fin Distress* | 7.249 | 7.229 | 6.994 |
| *Non-Fin Distress* | 4.810 | 5.412 | 5.372 |
| *Difference* | 2.439*** | 1.817*** | 1.622*** |
| ***LnRevenue*** | | | |
| *Fin Distress* | 16.995 | 16.966 | 16.947 |
| *Non-Fin Distress* | 17.121 | 17.130 | 17.134 |
| *Difference* | -0.126*** | -0.164*** | -0.187*** |
| ***Revenue development (%)*** | | | |
| *Fin Distress* | 0.066 | 0.074 | 0.095 |
| *Non-Fin Distress* | 0.087 | 0.088 | 0.080 |
| *Difference* | -0.021 | -0.014 | -0.015 |
| ***Financial Distress lag*** | | | |
| *Fin Distress* | 0.731 | 0.660 | 0.603 |
| *Non-Fin Distress* | 0.235 | 0.299 | 0.333 |
| *Difference* | 0.496*** | 0.361 | -0.270 |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

The Financial Distress lag variable is added to both tables, because it will be used in both regressions. The difference in the means of this variable is only significant at the first lag. In Table 8, absenteeism, staff turnover, and accounts receivable are significant at the 1%-level for all lags. Absenteeism and accounts receivable have a clear trend in their mean differences. Competition is only significant for lag 1 and 2 at the 1% and 5%-level. The annual statement is significant for one and two years before the event date at the 1%-level and book value over the purchase value is only significant at the first lag at the 5%-level. Just as in Table 5, Competition and book value over purchase value are not in line with the expectations.

Table 8: Lagged variables of the operational variables

| *Operational variables* | *Lag 1* | *Lag 2* | *Lag 3* |
|---|---|---|---|
| ***Absenteeism*** | | | |
| *Fin Distress* | 0.054 | 0.053 | 0.052 |
| *Non-Fin Distress* | 0.051 | 0.051 | 0.051 |
| *Difference* | 0.003*** | 0.002*** | 0.001*** |
| ***Staff Turnover*** | | | |
| *Fin Distress* | 0.144 | 0.144 | 0.140 |
| *Non-Fin Distress* | 0.131 | 0.130 | 0.130 |
| *Difference* | 0.013*** | 0.014*** | 0.010*** |
| ***Competition*** | | | |
| *Fin Distress* | 11506.69 | 11012.12 | 10550.03 |
| *Non-Fin Distress* | 9171.80 | 9477.15 | 9973.14 |
| *Difference* | 2334.89*** | 1534.97** | 576.89 |
| ***Book value/ Purchase value*** | | | |
| *Fin Distress* | 0.514 | 0.514 | 0.518 |
| *Non-Fin Distress* | 0.507 | 0.510 | 0.510 |
| *Difference* | 0.007* | 0.004 | 0.008 |
| ***Accounts receivable*** | | | |
| *Fin Distress* | 0.034 | 0.033 | 0.031 |
| *Non-Fin Distress* | 0.027 | 0.027 | 0.026 |
| *Difference* | 0.007*** | 0.006*** | 0.005*** |
| ***Annual Statement*** | | | |
| *Fin Distress* | 0.374 | 0.389 | 0.408 |
| *Non-Fin Distress* | 0.318 | 0.349 | 0.390 |
| *Difference* | 0.056*** | 0.040*** | 0.018 |
| ***Financial Distress lag*** | | | |
| *Fin Distress* | 0.731 | 0.660 | 0.603 |
| *Non-Fin Distress* | 0.235 | 0.299 | 0.333 |
| *Difference* | 0.496*** | 0.361 | -0.270 |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Most of the variables are significantly different for institutions in FD. RevenueD is the only variable that does not have a significant difference for any lag. The significant differences could increase the probability of a high prediction accuracy.

# 4. Methodology

The two main methods applied in this study are the Altman Z-model and the logistic regression model.

## 4.1. Correlation test

Before doing the regressions in Stata a correlation test needs to be done. When two independent variables are significantly correlated with one another, one of the variables will be removed from the regression. If both variables are held, even though they are highly correlated, there is multicollinearity in the regression, and this can lead to biased results. A variable is highly correlated when the correlation between these variables is at least 50%. The variable that will be removed depends on different criteria. The first criterium is based on the number of highly correlated variables a single variable has. For example, if variable A and variable B are highly correlated with one another and variable A is also highly correlated with another variable, variable A will be removed from the regression so that the least variables are lost. If both variables have the same amount of high correlations, the popularity of the variable in previous papers will play a role. If variable A and B are mentioned an equal amount of times in the papers discussed in the literature section, then the variable is chosen based on the significant difference between the mean of an HCI in FD, and not in FD. The variable with the better significance level will be used for the regression.

It is important to only keep one variable in the regression, because if both of the variables are used in the regression the predictive power will be weaker, and it will not be clear which of the two variables really affect the performance of the firm (Altman, 1968).

Lacher et al. (1995) mentions that it is crucial to choose variables that are not highly correlated to maximize discriminant power. The coefficients that are acquired by the MDA have enough discriminant power to classify a random institution into a certain group.

## 4.2. Altman model

This thesis consists of two groups for classifying an HCI. The dimensionality is obtained by G-1, where G stands for the number of groups. So, here the number of dimensionalities is equal to 1. The formula for classifying an HCI into a prespecified group has the form $Z = v_1x_1\ v_2x_2 + \ldots + v_nx_n$. This formula transforms the individual values to a single Z-value.

$v_1, v_2, \dots v_n$ = Discriminant coefficients

$x_1, x_2, \dots x_n$ = Independent variables

As mentioned in the literature, Altman his original model was not suitable for private firms. Therefore, the revised Altman discriminant function is used in this study. The formula for the revised Altman model is:

$Z\_i,t = 0.717x_{1\,(i,t-k)} + 0.847x_{2\,(i,t-k)} + 3.107x_{3\,(i,t-k)} + 0.420x_{4\,(i,t-k)} + 0.998x_{5\,(i,t-k)}$, for k= 1, 2 and 3

i = an HCI in the dataset

k = lag number

The coefficients, the variables and the Z-score are different in the revised model compared to the original Altman model. The variables for the revised Altman model are:

$x_1$ = Working capital/ Total Assets

$x_2$ = Operating results/ Total Assets

$x_3$ = EBIT/ Total Assets

$x_4$ = Bookvalue equity/ Total Liabilities

$x_5$ = Revenue/ Total Assets

The revised model is different for $x_4$. $X_4$ is equal to the book value of equity (BVETL) instead of the market value to equity (MVETL). As mentioned in the literature, Altman (1983) was trying to predict bankruptcies for private firms and these firms do not have market equity. The Z-score is calculated by taking the variables for one, two and three years before the eventdate and multiply them individually by the coefficients, so not all the lags together in the same formula. The multiplied outcomes are then added together. This will lead to one number which is classified as the Z-score. Different Z-scores have different interpretations. There are three Z-score zones to compare with the calculated Z-score. The Z-scores are acquired from the Altman paper in 1983. These Z-scores zones are:

- Z <= 1.23, this means that the health care institution is in the financial distress zone.
- 1.23 < Z < 2.9, this means that the health care institution is in the grey zone.
- Z >= 2.9, this means that the health care institution is in the safe zone.

If the institution is in the grey-zone that HCI will not be classified to a particular group. When all the Z-scores are obtained, and the institutions are placed in their predicted zones, the outcome is compared to the actual zone of the HCI at t = 0. The correct and incorrect classifications are discussed at the end of this section.

It is not possible to make an own prediction of the Altman model and compare it with the coefficients of the revised Altman model, because Altman did not explain the way he composed the model.

## 4.3. Logistic regression

A logistic regression model gives the coefficients, which could then be translated to a probability, that a certain institution belongs more to a prespecified category. The categories are FD and non-FD HCI. The dummy is 1 for HCI in FD and 0 otherwise. If the variable is more likely linked to a FD HCI the outcome of the multiplication (variable times coefficient) will be closer to 1 and vice versa. Although Ohlson (1980) wrote:" There is no apparent reason why 0.5 is an appropriate cut-off point, since it presumes implicitly that the loss function is symmetric across the two types of classification errors" the cut-off point of 0.5 will still be used. The cut-off point of 0.5 is the default value, from there it is easier to compare the threshold with other cut-off points and analyze the differences. The tables in the appendix, show the results for the cut-off points of 0.45 and 0.55. A cut-off point closer to 0, will correctly predict more HCI in FD than a cut-off point closer to 1.

A logistic regression is used when a linear regression is unable to explain a certain situation. The main problem with linear regressions is that the probabilities are not capped between 0 and 1. So in some situations, it is possible to have a negative probability or a probability higher than 100%. To get rid of this problem it is better to use a logistic regression. "The logit model has no restrictive assumptions for bankruptcy prediction." Kim and Gu (2006). The logistic formula of Kim and Gu (2006) is applied here.

The logistic regression model is:

$$Ln\left(\frac{Pi}{1-Pi}\right) = \beta_{0\,(i,\,t\text{-}k)} + \beta_1 X_{1(i,\,t\text{-}k)} + \beta_2 X_{2(i,\,t\text{-}k)} + \ldots + \beta_n X_{n\,(i,\,t\text{-}k)} \text{ for } k = 1, 2 \text{ and } 3.$$

i = a company in the dataset

k = the lag

$n$ = the number of the variable

The Altman variables for the logistic regression are:

$x_1$ = Working capital/ Total Assets

$x_2$ = Operating results/ Total Assets

$x_3$ = EBIT/ Total Assets

$x_4$ = Bookvalue equity/ Total Liabilities

$x_5$ = Revenue/ Total Assets

The financial variables for the logistic regression are (Finance Ideas):

$x_1$ = Current ratio

$x_2$ = Efficiency ratio

$x_3$ = Result ratio

$x_4$ = Net debt/EBITDA

$x_5$ = LnRevenue

$x_6$= Revenue development (%)

$x_7$ = Financial Distress lag

The operational variables for the logistic regression are (Finance Ideas):

$x_1$ = Absenteeism

$x_2$ = Staff turnover

$x_3$ = Competition

$x_4$ = Book value over Purchase value

$x_5$ = Accounts receivable

$x_6$ = Annual statement

$x_7$ = Financial Distress lag

The combined variables for the logistic regression are all the variables previously mentioned. So, the Altman, the financial and the operational variables, excluding the variables that are highly correlated with one another.

For every unit increase/decrease of variable $X_n$, ln $(\frac{Pi}{1-Pi})$ increases/decreases by $\beta_n$ times $X_n$, when $\beta_n$ is positive. When $\beta_n$ is negative, the outcome decreases/increases when $X_n$ increases/decreases. Only the significant coefficients have an effect on ln $(\frac{Pi}{1-Pi})$. All the $\beta_n$'s will be multiplied with their $X_n$'s and added together. This number is equal to the ln ($\frac{Pi}{1-Pi}$). It is possible to substitute ln ($\frac{Pi}{1-Pi}$) with y* to make it easier to go from the outcome of the formula to the probability.

$$y^* = \ln \left(\frac{Pi}{1-Pi}\right)$$

To go from y* to a probability, the formula below can be used:

$$Pi = \frac{\exp(y*)}{\exp(y*)+1} = \frac{\exp(b0+b1X1+b2X2+\cdots+bnXn)}{\exp(b0+b1X1+b2X2+\cdots+bnXn)+1}$$

The probabilities are calculated for all three lags and all companies.

## 4.4 Outcomes & Case

There are four different outcomes of the Altman and logistic test that will be stated below:

- When an HCI in FD is correctly classified, this is stated in the table as H1.
- When an HCI not in FD is correctly classified, this is stated in the table as H2.
- When an HCI in FD is classified as an HCI not in FD, this is considered a Type I error (T1E).
- When an HCI not in FD is classified as an HCI in FD, this is considered a Type II error (T2E).

The total hit score is the number of correctly classified firms, regardless if it was an HCI in financial distress or a healthy HCI. It is important to keep in mind that, there is a possibility an HCI can be assigned to the incorrect group. The table below gives extra insight in how the outcomes work.

Table 9A: Prediction accuracy model

| **Actual Group Membership (cut-off point)** | **Predicted Group Membership** | |
|---|---|---|
| **(Variable group) (correctly classified)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | **H1** (Number of HCI) (Percentage classified) | **T1E** (Number of HCI) (Percentage classified) |
| **Non-Financial Distress** | **T2E** (Number of HCI) (Percentage classified) | **H2** (Number of HCI) (Percentage classified) |

The percentage of correctly classified HCI needs to be significantly higher than 50%, in order to use the model for future predictions. Anything under 50% is really bad, because with simple guessing or a coin toss there is a better outcome. If the predictive power is not significantly different from 50%, it does not matter if someone uses this model or just does a coin toss in predicting FD, so I would not recommend the model. But, if the predictive power is significantly different from 50%, the model can be used to predict FD, because it is more likely for a model with a higher prediction accuracy to correctly predict the position of an HCI in the next period.

If the model has a bad prediction accuracy although the coefficients are all significant, I will randomly pick two institutions out of the data set that are in FD on t = 0 and two that are not in FD at t = 0. Afterwards, I will analyze them more thoroughly to better understand the outcome of the model. The table below will be filled in for the first, second and third lag variables. Remember that an institution that is classified in FD has a probability higher than 50%.

Table 9B: Case study

| | Institution 1 | Institution 2 | Institution 3 | Institution 4 |
|---|---|---|---|---|
| Actual | FD/Not in FD | FD/Not in FD | FD/Not in FD | FD/Not in FD |
| $X_1$ | % | % | % | % |
| $X_2$ | % | % | % | % |
| $X_n$ | % | % | % | % |
| Probability | % (FD/Not in FD) | % (FD/Not in FD) | % (FD/Not in FD) | % (FD/Not in FD) |
| Correct? | Yes/No | Yes/No | Yes/No | Yes/No |

The robustness of the results is checked by:

- Splitting the sample in two periods (2009-2013 and 2014-2018) and:
  - Compare the results of the Altman variables in the Altman model and the logistic regression with each other and the full sample.
  - Compare the results of the combined variables for both periods with the full sample.

# 5. Results

In this section, the correlation for the Altman, financial, operational and the combined variables will be specified first. After the correlation tests, the results of the Altman variables with the Altman model, followed by the results of the Altman variables with the logistic regression will be compared to each other. Finally, the results of the financial, operational, and the combined variables with the logistic regression are extensively discussed. The logistic regressions with the cut-off points at 0.45 and 0.55 can be found in the Appendix.

## 5.1. Correlation

### 5.1.1. Correlation Altman variables

Table 10 shows the correlation of the Altman variables between one other. None of the Altman variables are correlated with the dependent variable. The working capital over total assets (WCTA) is highly correlated with book value of equity over total liabilities (BVETL) on all lags, with a correlation score of respectively 0.654, 0.654 and 0.649. The operational results over total assets (ORTA) is highly correlated with the EBIT over total assets (EBITTA) for all lags, and has correlation values of respectively 0.725, 0.721 and 0.712. BVETL will be left out of the regression, because WCTA is mentioned in more papers, and is also from the original model. ORTA will be left out of the regression because EBITTA is from the original model and ORTA is a proxy of retained earnings over total assets. When these two variables are removed, there are no highly correlated variables left in the regression.

Table 10: Correlation table Altman variables

| *First Lag Altman Variables* | *FD* | *WCTA* | *ORTA* | *EBITTA* | *BVETL* | *RTA* |
|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | |
| *WCTA* | -0.197 | 1.000 | | | | |
| *ORTA* | -0.239 | 0.330 | 1.000 | | | |
| *EBITTA* | -0.218 | 0.007 | **0.725** | 1.000 | | |
| *BVETL* | -0.193 | **0.654** | 0.237 | 0.031 | 1.000 | |
| *RTA* | 0.058 | 0.050 | 0.126 | 0.014 | -0.002 | 1.000 |
| *Second Lag Altman Variables* | FD | WCTA | ORTA | EBITTA | BVETL | RTA |
| *FD* | 1.000 | | | | | |
| *WCTA* | -0.141 | 1.000 | | | | |
| *ORTA* | -0.159 | 0.332 | 1.000 | | | |
| *EBITTA* | -0.152 | 0.006 | **0.721** | 1.000 | | |
| *BVETL* | -0.154 | **0.654** | 0.252 | 0.041 | 1.000 | |
| *RTA* | 0.060 | 0.068 | 0.145 | 0.033 | 0.016 | 1.000 |
| *Third Lag Altman Variables* | FD | WCTA | ORTA | EBITTA | BVETL | RTA |
| *FD* | 1.000 | | | | | |
| *WCTA* | -0.102 | 1.000 | | | | |
| *ORTA* | -0.137 | 0.339 | 1.000 | | | |
| *EBITTA* | -0.141 | 0.015 | **0.712** | 1.000 | | |
| *BVETL* | -0.133 | **0.649** | 0.266 | 0.053 | 1.000 | |
| *RTA* | 0.068 | 0.008 | 0.167 | 0.048 | 0.034 | 1.000 |

### 5.1.2. Correlation Financial variables

In Table 11, the correlation between the financial variables are presented. Just as the Altman variables, the financial variables are not highly correlated with the dependent variable. No independent variable is highly correlated with another independent variable, which means that all the variables will be added to the regression.

Table 11: Correlation table Financial variables

| *First Lag Financial Variables* | *FD* | *CR* | *ER* | *RR* | *DEBITDA* | *LnRevenue* | *RevenueD (%)* |
|---|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | | |
| *CR* | -0.151 | 1.000 | | | | | |
| *ER* | 0.081 | -0.178 | 1.000 | | | | |
| *RR* | -0.279 | 0.338 | -0.465 | 1.000 | | | |
| *DEBITDA* | 0.111 | -0.142 | 0.003 | -0.102 | 1.000 | | |
| *LnRevenue* | -0.059 | -0.334 | 0.164 | -0.173 | 0.065 | 1.000 | |
| *RevenueD (%)* | -0.012 | -0.009 | -0.033 | 0.038 | 0.003 | -0.032 | 1.000 |
| *Second Lag Financial Variables* | FD | CR | ER | RR | DEBITDA | LnRevenue | RevenueD (%) |
| *FD* | 1.000 | | | | | | |
| *CR* | -0.120 | 1.000 | | | | | |
| *ER* | 0.070 | -0.186 | 1.000 | | | | |
| *RR* | -0.209 | 0.345 | -0.461 | 1.000 | | | |
| *DEBITDA* | 0.080 | -0.131 | 0.006 | -0.1064 | 1.000 | | |
| *LnRevenue* | -0.063 | -0.354 | 0.179 | -0.193 | 0.074 | 1.000 | |
| *RevenueD (%)* | -0.073 | -0.010 | -0.030 | 0.030 | 0.004 | -0.031 | 1.000 |
| *Third Lag Financial Variables* | FD | CR | ER | RR | DEBITDA | LnRevenue | RevenueD (%) |
| *FD* | 1.000 | | | | | | |
| *CR* | -0.088 | 1.000 | | | | | |
| *ER* | 0.057 | -0.199 | 1.000 | | | | |
| *RR* | -0.181 | 0.357 | -0.462 | 1.000 | | | |
| *DEBITDA* | 0.074 | -0.136 | 0.002 | -0.117 | 1.000 | | |
| *LnRevenue* | -0.068 | -0.370 | 0.206 | -0.207 | 0.072 | 1.000 | |
| *RevenueD (%)* | -0.007 | -0.010 | -0.028 | 0.024 | 0.002 | -0.030 | 1.000 |

### 5.1.3. Correlation Operational variables

Table 12 states the correlation between the operational variables. None of the variables are highly correlated with the dependent variable, and no variable is highly correlated with another. All the variables will be used for the regression.

| *First Lag Operational Variables* | *FD* | *Absent* | *ST* | *Comp* | *BVPV* | *OnderW* | *AS* | *FD lag* |
|---|---|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | | | |
| *Absent* | 0.073 | 1.000 | | | | | | |
| *ST* | 0.048 | 0.094 | 1.000 | | | | | |
| *Comp* | 0.047 | -0.163 | -0.063 | 1.000 | | | | |
| *BVPV* | 0.021 | 0.029 | -0.005 | 0.069 | 1.000 | | | |
| *OnderW* | 0.073 | -0.147 | 0.042 | 0.077 | -0.112 | 1.000 | | |
| *AS* | 0.060 | -0.069 | 0.040 | -0.113 | -0.259 | 0.084 | 1.000 | |
| *FD lag* | 0.494 | 0.082 | 0.044 | 0.091 | 0.025 | 0.079 | 0.067 | 1.000 |
| *Second Lag Operational Variables* | FD | Absent | ST | Comp | BVPV | OnderW | AS | FD lag |
| *FD* | 1.000 | | | | | | | |
| *Absent* | 0.049 | 1.000 | | | | | | |
| *ST* | 0.050 | 0.093 | 1.000 | | | | | |
| *Comp* | 0.029 | -0.158 | -0.060 | 1.000 | | | | |
| *BVPV* | 0.011 | 0.034 | -0.000 | 0.075 | 1.000 | | | |
| *OnderW* | 0.056 | -0.152 | 0.034 | 0.105 | -0.108 | 1.000 | | |
| *AS* | 0.043 | -0.062 | 0.045 | -0.136 | -0.296 | 0.095 | 1.000 | |
| *FD lag* | 0.356 | 0.075 | 0.045 | 0.113 | | | | 1.000 |
| *Third Lag Operational Variables* | FD | Absent | ST | Comp | BVPV | OnderW | AS | FD lag |
| *FD* | 1.000 | | | | | | | |
| *Absent* | 0.042 | 1.000 | | | | | | |
| *ST* | 0.034 | 0.095 | 1.000 | | | | | |
| *Comp* | 0.008 | -0.154 | -0.056 | 1.000 | | | | |
| *BVPV* | 0.021 | 0.035 | -0.001 | 0.082 | 1.000 | | | |
| *OnderW* | 0.054 | -0.151 | 0.028 | 0.135 | -0.113 | 1.000 | | |
| *AS* | 0.021 | -0.057 | 0.048 | 0.153 | -0.323 | 0.112 | 1.000 | |
| *FD lag* | 0.266 | 0.059 | 0.041 | 0.141 | 0.023 | 0.111 | 0.059 | 1.000 |

5.1.4. Correlation Combined variables

Table 13A, B and C show the correlation of all the variables that were mentioned in the sections above. The variables that are highly correlated for all lags are, WCTA with BVETL and CR ORTA with EBITTA and RR, EBITTA with RR, and BVEIL with CR. The correlation between WCTA/BVETL*, and WCTA/CR are for the one, two and three-year lag respectively 0.660,0.661, 0.657 and 0.835, 0.838 and 0.840. WCTA is the most mentioned variable in earlier papers, and therefore, more important compared to BVETL and CR, which mean that these two variables are removed from the combined regression. EBITTA is correlated with ORTA and RR. The correlation between EBITTA/ORTA, and EBITTA/RR for the one, two and three-year lag are respectively 0.719, 0.714, 0.701, and 0.853, 0.850 and 0.841. ORTA and RR are not mentioned in any paper in this study's literature section, ORTA is used as a proxy for retained earnings and RR is used by Finance Ideas. EBITTA is mentioned in three Altman papers so this will be the variable used in the regression and therefore, ORTA and RR will be dropped. In conclusion, BVETL, CR, ORTA, and RR are removed from the regression.

* WCTA/BVETL is the correlation between these two variables. Not one divided by the other. Same holds for X/Y. Where X and Y are variables.

Table 13A: Correlation one year lagged combined variables

a = Ln Revenue b = Absenteeism c = Competition d = Accounts receivable

| | *FD* | *WCTA* | *ORTA* | *EBITTA* | *BVETL* | *RTA* | *CR* | *ER* | *RR* | *DEBITDA* | *Ln Rev*[a] | *RevenueD* | *Absent*[b] | *ST* | *Comp*[c] | *BVPV* | *Acc rec*[d] | *AS* | *FD lag* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | | | | | | | | | | | | | | |
| *WCTA* | -0.174 | 1.000 | | | | | | | | | | | | | | | | | |
| *ORTA* | -0.228 | 0.317 | 1.000 | | | | | | | | | | | | | | | | |
| *EBITTA* | -0.214 | -0.020 | **0.719** | 1.000 | | | | | | | | | | | | | | | |
| *BVETL* | -0.178 | **0.660** | 0.230 | 0.018 | 1.000 | | | | | | | | | | | | | | |
| *RTA* | 0.077 | 0.043 | 0.139 | 0.019 | -0.002 | 1.000 | | | | | | | | | | | | | |
| *CR* | -0.149 | **0.835** | 0.227 | -0.037 | **0.795** | -0.062 | 1.000 | | | | | | | | | | | | |
| *ER* | 0.080 | -0.121 | -0.342 | -0.373 | -0.135 | 0.224 | -0.179 | 1.000 | | | | | | | | | | | |
| *RR* | -0.279 | 0.345 | **0.853** | **0.585** | 0.297 | -0.080 | 0.339 | -0.463 | 1.000 | | | | | | | | | | |
| *DEBITDA* | 0.111 | -0.159 | -0.071 | -0.0814 | -0.178 | -0.089 | -0.143 | 0.003 | -0.101 | 1.000 | | | | | | | | | |
| *Ln Rev*[a] | -0.055 | -0.358 | -0.196 | 0.058 | -0.269 | -0.346 | -0.336 | 0.161 | -0.173 | 0.065 | 1.000 | | | | | | | | |
| *RevenueD* | -0.012 | 0.005 | 0.052 | 0.038 | -0.013 | 0.077 | -0.009 | -0.032 | 0.036 | 0.003 | -0.031 | 1.000 | | | | | | | |
| *Absent*[b] | 0.079 | -0.070 | -0.217 | -0.129 | -0.057 | -0.064 | -0.070 | 0.138 | -0.220 | 0.0008 | 0.104 | -0.052 | 1.000 | | | | | | |
| *ST* | 0.055 | 0.001 | 0.001 | -0.077 | -0.020 | 0.149 | -0.018 | 0.092 | -0.043 | 0.024 | -0.073 | -0.004 | 0.081 | 1.000 | | | | | |
| *Comp*[c] | 0.025 | -0.152 | -0.058 | 0.104 | -0.148 | -0.201 | -0.137 | -0.031 | -0.016 | 0.020 | 0.411 | -0.013 | -0.180 | -0.075 | 1.000 | | | | |
| *BVPV* | 0.009 | -0.198 | -0.048 | 0.005 | -0.109 | -0.224 | -0.132 | -0.065 | -0.003 | 0.038 | 0.188 | -0.002 | -0.001 | -0.039 | 0.031 | 1.000 | | | |
| *Acc rec*[d] | 0.072 | 0.051 | 0.058 | -0.068 | -0.059 | 0.089 | -0.031 | 0.078 | 0.023 | 0.029 | -0.068 | 0.020 | -0.140 | 0.044 | 0.066 | -0.113 | 1.000 | | |
| *AS* | 0.071 | -0.036 | 0.045 | -0.026 | -0.066 | 0.112 | -0.052 | -0.009 | 0.017 | 0.051 | -0.216 | 0.023 | -0.059 | 0.058 | -0.057 | -0.176 | 0.088 | 1.000 | |
| *FD lag* | 0.484 | -0.300 | -0.384 | -0.360 | -0.267 | 0.069 | -0.260 | 0.138 | -0.434 | 0.153 | -0.024 | -0.019 | 0.091 | 0.048 | 0.070 | 0.037 | 0.074 | 0.071 | 1.000 |

Table 13: Correlation one year lagged combined variables

a = Ln Revenue b = Absenteeism c = Competition d = Accounts receivable

| | *FD* | *WCTA* | *ORTA* | *EBITTA* | *BVETL* | *RTA* | *CR* | *ER* | *RR* | *DEBITDA* | *Ln Rev*[a] | *RevenueD* | *Absent*[b] | *ST* | *Comp*[c] | *BVPV* | *Acc rec*[d] | *AS* | *FD lag* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | | | | | | | | | | | | | | |
| *WCTA* | -0.126 | 1.000 | | | | | | | | | | | | | | | | | |
| *ORTA* | -0.165 | 0.315 | 1.000 | | | | | | | | | | | | | | | | |
| *EBITTA* | -0.159 | -0.029 | **0.714** | 1.000 | | | | | | | | | | | | | | | |
| *BVETL* | -0.150 | **0.661** | 0.246 | 0.027 | 1.000 | | | | | | | | | | | | | | |
| *RTA* | 0.080 | 0.065 | 0.163 | 0.041 | 0.018 | 1.000 | | | | | | | | | | | | | |
| *CR* | -0.117 | **0.838** | 0.236 | -0.036 | **0.792** | -0.044 | 1.000 | | | | | | | | | | | | |
| *ER* | 0.068 | -0.127 | -0.337 | -0.360 | -0.141 | 0.212 | -0.187 | 1.000 | | | | | | | | | | | |
| *RR* | -0.208 | 0.344 | **0.850** | **0.575** | 0.312 | -0.071 | 0.346 | -0.459 | 1.000 | | | | | | | | | | |
| *DEBITDA* | 0.079 | -0.151 | -0.078 | -0.097 | -0.174 | -0.095 | -0.131 | 0.005 | -0.105 | 1.000 | | | | | | | | | |
| *Ln Rev*[a] | -0.067 | -0.386 | -0.219 | 0.050 | -0.280 | -0.348 | -0.355 | 0.176 | -0.194 | 0.074 | 1.000 | | | | | | | | |
| *RevenueD* | -0.008 | 0.002 | 0.044 | 0.031 | -0.014 | 0.080 | -0.010 | -0.028 | 0.028 | 0.004 | -0.029 | 1.000 | | | | | | | |
| *Absent*[b] | 0.044 | -0.067 | -0.227 | -0.130 | -0.058 | -0.077 | -0.071 | 0.136 | -0.227 | 0.025 | 0.102 | -0.052 | 1.000 | | | | | | |
| *ST* | 0.054 | 0.008 | 0.006 | -0.078 | -0.015 | 0.145 | -0.014 | 0.101 | -0.041 | 0.027 | -0.067 | -0.004 | 0.080 | 1.000 | | | | | |
| *Comp*[c] | 0.007 | -0.160 | -0.067 | 0.103 | -0.151 | -0.203 | -0.142 | -0.012 | -0.021 | 0.021 | 0.406 | -0.012 | -0.176 | -0.074 | 1.000 | | | | |
| *BVPV* | 0.004 | -0.189 | -0.060 | -0.006 | -0.099 | -0.216 | -0.128 | -0.058 | -0.010 | 0.026 | 0.197 | -0.004 | 0.005 | -0.035 | 0.033 | 1.000 | | | |
| *Acc rec*[d] | 0.060 | 0.037 | 0.056 | -0.054 | -0.061 | 0.067 | -0.034 | 0.080 | 0.025 | 0.035 | -0.052 | 0.018 | -0.146 | 0.035 | 0.098 | -0.108 | 1.000 | | |
| *AS* | 0.047 | -0.011 | 0.055 | -0.042 | -0.051 | 0.125 | -0.030 | -0.005 | 0.024 | 0.031 | -0.236 | 0.019 | -0.050 | 0.064 | -0.077 | -0.212 | 0.100 | 1.000 | |
| *FD lag* | 0.348 | -0.301 | -0.376 | 0.352 | -0.272 | 0.042 | -0.259 | 0.124 | -0.424 | 0.151 | -0.008 | -0.017 | 0.081 | 0.049 | 0.092 | 0.037 | 0.082 | 0.064 | 1.000 |

Table 13: Correlation one year lagged combined variables

a = Ln Revenue b = Absenteeism c = Competition d = Accounts receivable

| | *FD* | *WCTA* | *ORTA* | *EBITTA* | *BVETL* | *RTA* | *CR* | *ER* | *RR* | *DEBITDA* | *Ln Rev*[a] | *RevenueD* | *Absent*[b] | *ST* | *Comp*[c] | *BVPV* | *Acc rec*[d] | *AS* | *FD lag* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *FD* | 1.000 | | | | | | | | | | | | | | | | | | |
| *WCTA* | -0.085 | 1.000 | | | | | | | | | | | | | | | | | |
| *ORTA* | -0.128 | 0.319 | 1.000 | | | | | | | | | | | | | | | | |
| *EBITTA* | -0.137 | -0.026 | **0.701** | 1.000 | | | | | | | | | | | | | | | |
| *BVETL* | -0.120 | **0.657** | 0.262 | 0.039 | 1.000 | | | | | | | | | | | | | | |
| *RTA* | 0.092 | 0.079 | 0.191 | **0.061** | 0.038 | 1.000 | | | | | | | | | | | | | |
| *CR* | -0.086 | **0.840** | 0.243 | -0.0315 | **0.788** | -0.029 | 1.000 | | | | | | | | | | | | |
| *ER* | 0.058 | -0.141 | -0.331 | 0.348 | -0.152 | 0.210 | -0.198 | 1.000 | | | | | | | | | | | |
| *RR* | -0.183 | 0.354 | **0.841** | **0.547** | 0.327 | -0.074 | 0.358 | -0.461 | 1.000 | | | | | | | | | | |
| *DEBITDA* | 0.073 | -0.152 | -0.090 | -0.121 | -0.176 | -0.104 | -0.136 | 0.002 | -0.115 | 1.000 | | | | | | | | | |
| *Ln Rev*[a] | -0.072 | -0.405 | -0.243 | 0.037 | -0.287 | -0.350 | -0.371 | 0.204 | -0.210 | 0.072 | 1.000 | | | | | | | | |
| *RevenueD* | -0.007 | 0.003 | 0.038 | 0.027 | -0.014 | 0.081 | -0.010 | -0.026 | 0.023 | 0.002 | -0.029 | 1.000 | | | | | | | |
| *Absent*[b] | 0.048 | -0.062 | -0.219 | -0.119 | -0.052 | -0.068 | -0.065 | 0.126 | -0.211 | 0.032 | 0.092 | -0.050 | 1.000 | | | | | | |
| *ST* | 0.034 | 0.013 | 0.015 | -0.077 | -0.008 | 0.154 | 0.010 | 0.115 | -0.034 | 0.050 | -0.065 | -0.004 | 0.085 | 1.000 | | | | | |
| *Comp*[c] | -0.025 | -0.163 | -0.084 | 0.090 | -0.154 | -0.206 | -0.144 | 0.023 | -0.038 | 0.024 | 0.400 | -0.012 | -0.174 | -0.071 | 1.000 | | | | |
| *BVPV* | 0.022 | -0.179 | -0.077 | -0.020 | -0.093 | -0.197 | -0.124 | -0.051 | -0.021 | 0.009 | 0.202 | -0.007 | 0.005 | -0.040 | 0.035 | 1.000 | | | |
| *Acc rec*[d] | 0.043 | 0.019 | 0.043 | -0.054 | -0.080 | 0.044 | 0.051 | 0.072 | 0.010 | 0.029 | -0.030 | 0.018 | -0.143 | 0.029 | 0.134 | -0.106 | 1.000 | | |
| *AS* | 0.034 | 0.015 | 0.065 | -0.057 | -0.042 | 0.123 | 0.011 | 0.002 | 0.024 | 0.040 | -0.255 | 0.016 | -0.044 | 0.068 | -0.087 | -0.238 | 0.120 | 1.000 | |
| *FD lag* | 0.255 | -0.303 | -0.362 | -0.335 | -0.279 | 0.029 | 0.254 | 0.108 | -0.407 | 0.161 | 0.008 | -0.016 | 0.067 | 0.048 | 0.122 | 0.039 | 0.108 | 0.066 | 1.000 |

## 5.2. Results Altman model and logistic regression full sample

### 5.2.1. Altman model full sample

Table 14 shows the results of the first test. The Altman test has a good prediction rate for institutions in FD. Variables that are further from the date of the event have a less accurate prediction rate. This means that more recent variables lead to a better prediction accuracy than variables that are more in the past: this finding is in line with the theory. The highest prediction accuracy is 80.52% for correctly predicting firms in FD one year before the event date (lag1).

Table 14: Results Altman Model.

| **Actual Group Membership** | **Predicted Group Membership** | |
|---|---|---|
| **Lag 1 Altman model (64.62%)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | 496 (80.52%) | 881 (39.81%) |
| **Non-Financial Distress** | 120 (19.48%) | 1332 (60.19%) |
| **Lag 2 Altman model (62.05%)** | | |
| **Financial Distress** | 425 (75.35%) | 803 (41,87%) |
| **Non-Financial Distress** | 139 (24,65%) | 1115 (58,13%) |
| **Lag 3 Altman model (60.96%)** | | |
| **Financial Distress** | 350 (71,28%) | 693 (42,13%) |
| **Non-Financial Distress** | 141 (28,72%) | 952 (57,87%) |

In 80.52% of the cases, the Altman model is able to accurately predict, one year beforehand, if an HCI actually is in FD at t = 0. It is also interesting to note that the model is better in predicting HCI in FD than HCI not in FD. The results are weaker than the results of previous Altman papers. In earlier studies, the prediction accuracy for one, two and three-year lags, were all around 80% or more. One year before the event date, the model had prediction accuracies of 85% and up. The results for accurately predicting FD, two and three years before the event date, were 75.35% and 71.28%. For non-FD HCI the prediction accuracy for one, two and three years before the event date were respectively 60.19%, 58.13% and 57.87%. The total percentage of correctly classified HCI of the Altman model for the one, two, and three-year lag(s) are 64.62%,

62.05% and 60.96%. Only the HCI that were in the distress and safe zone were classified. The number of HCI for the Altman model are lower than for the logistic regressions, because only the observations in the FD and safe zone were classified, not in the grey zone.

5.2.2.1. Altman logistic regression full sample

It is important to compare the results of the Altman model with the result of the logistic regression, to conclude which model is better in predicting. As mentioned in the correlation section of the Altman variables, the ORTA and BVETL were removed from the regression. The coefficients in Table 15 of WCTA, EBITTA and RTA for the first lag are respectively -1.849, -5.196 and 0.163. If WCTA and EBITTA increase, this will lead to a

Table 15: Coefficients Altman variables

| *Altman variables* | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Working Capital/ Total Assets* | -1.849 | 0.000*** | -1.243 | 0.000*** | -0.899 | 0.000*** |
| *EBIT/ Total Assets* | -5.196 | 0.000*** | -3.286 | 0.000*** | -2.988 | 0.000*** |
| *Revenue/ Total Assets* | 0.163 | 0.000*** | 0.154 | 0.000*** | 0.167 | 0.000*** |
| *Constant* | 0.565 | 0.000*** | 0.169 | 0.012*** | 0.034 | 0.641 |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

lower ln $(\frac{Pi}{1-Pi})$, and eventually a number closer to 0 than to 1. This indicates that an HCI with a higher WCTA and higher EBITTA is more likely to be a healthy institution. When RTA increases, the ln $(\frac{Pi}{1-Pi})$ also increases, and the institution will more likely be in FD. Just like Table 6, the RTA is positive which is not according to the expectation. The strength of the coefficients of the WCTA and the EBITTA decrease when the lag increases. This is in line with the expectation. Variables that are more in the past should have a weaker effect than variables that are closer to the date of the event. For RTA there is not a clear trend.

The results of Table 16 are very different from the results of Table 14. The results of Table 14 have a high prediction ratio for HCI in FD, contradicting the findings of Table 16, were the prediction rate is higher for institutions not in financial distress. Another notable point is that the prediction accuracy increases, when the lag increases for institutions not in FD. This means that variables further in the past, are better predictors than variables closer to the date of the event. Although this is not in line with the expectations, earlier research did not elaborate on this, when they had the same findings. Both models are okay in predicting the status of the institution with past variables, but they are not good predictors for the same part. A possible explanation could be that, only the variables that were clearly identified to a certain zone (distress/safe), were used for the prediction accuracy of the Altman model, whereas for the logistic regression, all HCI are used.

Table 16: Prediction accuracy Altman variables logistic regression

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | |
|---|---|---|
| **Lag 1 Altman model (64.10%)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | 1300 (44.35%) | 764 (20.43%) |
| **Non-Financial Distress** | 2931 (55.65%) | 2976 (79.57%) |
| **Lag 2 Altman model (60.29%)** | | |
| **Financial Distress** | 678 (26.99%) | 479 (14.46%) |
| **Non-Financial Distress** | 1834 (73.01%) | 2834 (85.54%) |
| **Lag 3 Altman model (60.21%)** | | |
| **Financial Distress** | 402 (19.10%) | 281 (9.75%) |
| **Non-Financial Distress** | 1703 (80.90%) | 2600 (90.25%) |

5.2.2.2. Case

As mentioned above, Table 16 is bad in predicting HCI in financial distress one, two and three years beforehand, and is increasing in prediction accuracy for institutions not in FD. In this paragraph I will take four institutions out of the data set and analyze them more thoroughly. I want to track down why the prediction accuracy increases, when the lag increases, for institutions not in FD. The status of the institutions are: “no FD” on t = 0. The institutions that are chosen from the dataset are: Altra, Proteion Thuis, Regionale Stichting Zorgcentra De Kempen, and Stichting Baalderborg Groep. The only criterium is that the first lag has predict that the HCI is in FD, before further analyzing the institution. If the second and third lag predict that the HCI is not in FD, after the first lag showed that it was in FD, this finding is in line with the pattern of table 16.

The formula for calculating the probability for FD in the first lag is:

$$Pi = \frac{\exp(0.565-1.849X1-5.196X2+0.163X3)}{\exp(0.565-1.849X1-5.196X2+0.163X3)+1}$$

Table 17A: Probability one-year lag

| | *Altra* | *Proteion Thuis* | *Regionale Stichting Zorgcentra De Kempen* | *Stichting Baalderborg Groep* |
|---|---|---|---|---|
| *Predicted* | 57.78% (FD) | 57.42% (FD) | 53.85% (FD) | 53.61% (FD) |
| *WCTA* | 26.05% | 0.75% | 7.25% | -7.71% |
| *EBITTA* | -2.37% | 9.15% | 9.58% | 14.84% |
| *RTA* | 216.98% | 137.12% | 135.55% | 127.61% |
| *Actual* | No FD | No FD | No FD | No FD |
| *Correct?* | No | No | No | No |

The formula for calculating the probability for FD in the second lag is:

$$Pi = \frac{\exp(0.169-1.243X1-3.286X2+0.154X3)}{\exp(0.169-1.243X1-3.286X2+0.154X3)+1}$$

Table 17B: Probability two-year lag

| | *Altra* | *Proteion Thuis* | *Regionale Stichting Zorgcentra De Kempen* | *Stichting Baalderborg Groep* |
|---|---|---|---|---|
| *Predicted* | 45.84% (No FD) | 50.43% (FD) | 46.37% (No FD) | 59.56% (FD) |
| *WCTA* | 28.35% | 7.94% | 9.21% | -6.89% |
| *EBITTA* | 8.13% | 7.41% | 12.07% | 14.39% |
| *RTA* | 184.30% | 123.60% | 127.64% | 135.97% |
| *Actual* | No FD | No FD | No FD | No FD |
| *Correct?* | Yes | No | Yes | No |

The formula for calculating the probability for FD in the third lag is:

$$Pi = \frac{\exp(0.034-0.899X1-2.998X2+0.167X3)}{\exp(0.034-0.899X1-2.998X2+0.167X3)+1}$$

Table 17C: Probability three-year lag

| | *Altra* | *Proteion Thuis* | *Regionale Stichting Zorgcentra De Kempen* | *Stichting Baalderborg Groep* |
|---|---|---|---|---|
| *Predicted* | 0.298% (No FD) | 43.48% (No FD) | 45.72% (No FD) | 44.23% (No FD) |
| *WCTA* | 23.28% | 13.59% | 6.70% | -5.32% |
| *EBITTA* | 35.14% | 12.17% | 12.10% | 17.93% |
| *RTA* | 224.39% | 114.31% | 130.05% | 133.98% |
| *Actual* | No FD | No FD | No FD | No FD |
| *Correct?* | Yes | Yes | Yes | Yes |

The analyses above tries to describe the situation of Table 16. The tables only show four analyses of more than 800 institutions but there are some interesting findings that are in line with Table 16. Table 17A shows that the variables in the first lag predict the HCI in FD, while the HCI is not in FD. When looking at table 17B, two of the four HCI are correctly classified, and table 17C classifies all HCI correctly. This is exactly what happens in table 16. Whenever the lag increases, the prediction accuracy for correctly predicting HCI not in FD also increases.

### 5.2.3. Financial variables

The third test is the logistic regression with the financial variables. All the variables are significant except for the efficiency ratio and the revenueD.

The signs for the current ratio (CR), the result ratio (RR), the LnRevenue and the Financial Distress lagged variable are negative, which indicates that institutions with a higher value for these variables, will less likely be in FD. If the net debt over EBITDA is higher the probability that the institution will be in financial distress will increase. There is no clear trend in the strength of the coefficients. According to the theory, the strength of the coefficients should increase if the coefficients are closer to the event date, but that is only the case for the Financial Distress lagged variable.

Table 18: Coefficients financial variables

| *Financial variables* | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Current ratio* | -0.069 | 0.016** | -0.089 | 0.002*** | -0.064 | 0.038** |
| *Efficiency ratio* | -0.643 | 0.133 | 0.086 | 0.841 | 0.026 | 0.956 |
| *Result ratio* | -5.918 | 0.000*** | -4.006 | 0.000*** | -5.065 | 0.000*** |
| *Debt/EBITDA* | 0.009 | 0.003*** | 0.006 | 0.045** | 0.007 | 0.039** |
| *LnRevenue* | -0.132 | 0.000*** | -0.146 | 0.000*** | -0.155 | 0.000*** |
| *RevenueD (%)* | -0.011 | 0.783 | -0.008 | 0.826 | -0.012 | 0.751 |
| *Fin Distress lag* | 1.854 | 0.000*** | 1.258 | 0.000*** | 0.834 | 0.000*** |
| *Constant* | 1.813 | 0.000*** | 1.755 | 0.001*** | 2.170 | 0.000*** |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

The percentage of correctly classified institutions, in Table 19, is good for one year prior to the date of the event. The correctly classified percentage for HCI one, two and three years before the event date were 74.63%, 68.21% and 64.00%. The 74.63% is a good accuracy for predicting the status of an institution, because it is a lot higher than 50%. This high accuracy indicates that it is better to use this model with variables one year before the event date than to take a guess at the outcome.

The accuracy for predicting institutions in financial distress, for one two and three years before the event, are respectively, 71.22%, 63.26% and 49.80%. The prediction accuracies for institutions not in financial distress are 77.20%, 71.82% and 74.53%. The financial variables have a higher percentage of correctly classified institutions compared to the results of the Altman variables. It is noteworthy that the prediction accuracy for non-FD HCI does not have a clear trend. I would only recommend using the model for the first lag, because the prediction accuracies for FD and non-FD are both significantly higher than 50%.

Table 19: Prediction accuracy financial variables

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | |
|---|---|---|
| **Lag 1 Financial variables (74.63%)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | 1787 (71.22%) | 755 (22.80%) |
| **Non-Financial Distress** | 722 (28.78%) | 2577 (77.20%) |
| **Lag 2 Financial variables (68.21%)** | | |
| **Financial Distress** | 1331 (63.26%) | 811 (28.18%) |
| **Non-Financial Distress** | 773 (36.74%) | 2067 (71.82%) |
| **Lag 3 Financial variables (64.00%)** | | |
| **Financial Distress** | 884 (49.80%) | 610 (25.47%) |
| **Non-Financial Distress** | 891 (50.20%) | 1785 (74.53%) |

### 5.2.4. Operational variables

The signs of the coefficients in Table 20 for Absenteeism, Staff Turnover, Accounts receivable, the Annual statement and the Financial Distress lag are in line with the expectation. When Absenteeism and Staff turnover increase, and the Annual statement and the Financial Distress lag is equal to 1, the probability that an HCI will be classified as an institution in FD will increase. The sign of the coefficient for Book value over Purchase value (BVPV) is not as expected. A low BVPV means that the institution lacked on the renovations. In theory HCI that lack on renovating their buildings, are HCI that have less capital. HCI that have less liquidity, are more likely to be in Financial distress. According to the expectation, the sign of the BVPV should be negative. A higher BVPV should lead to a lower outcome of the regression and a lower probability for FD. The competition coefficients are not significant for any lag.

It's also noteworthy that only the Absenteeism, Annual Statement and Financial Distress lag have a clear trend over time. The coefficients closer to the date of the event are higher than the coefficients that are more in the past except for Staff turnover, BVPV and accounts receivable.

Table 20: Coefficients Operational variables

| *Operational variables* | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Absenteeism* | 6.137 | 0.000*** | 3.135 | 0.063* | 3.076 | 0.083* |
| *Staff turnover* | 0.390 | 0.062* | 0.605 | 0.017** | 0.264 | 0.202 |
| *Competition (x100,000)* | 0.107 | 0.373 | -0.059 | 0.626 | 0.264 | 0.472 |
| *Book value/ Purchase value* | 0.270 | 0.086* | 0.136 | 0.382 | 0.238 | 0.026*** |
| *Accounts receivable* | 2.231 | 0.000*** | 1.416 | 0.027** | 1.662 | 0.000*** |
| *Annual statement* | 0.183 | 0.004*** | 0.107 | 0.091* | 0.017 | 0.071* |
| *FD lag* | 2.135 | 0.000*** | 1.477 | 0.000*** | 1.095 | 0.000*** |
| *Constant* | -1.922 | 0.000*** | -1.353 | 0.000*** | -1.166 | 0.000*** |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Table 21 shows the prediction accuracy of the operational variables at the 0.5 cut-off point. There is a clear trend for the prediction accuracy. The further the variables are in the past, the lower the prediction accuracy is. There is still a lot of time between the classification of an HCI, three years before the event date, and on t = 0 that can affect the outcome of the prediction. That is why the accuracy is lower for variables further from the event date, according to the theory. The highest percentage of correctly classified HCI is 75% for the first lag. This means that in 75% of the cases, a random HCI could be placed in the right category. The operational variables are better in classifying institutions not in FD than institutions in FD. The prediction accuracy for FD HCI one, two and three years before the event date are 73.04%, 65.75% and 58.45% and for HCI not in FD these are 76.35%, 70.05% and 67.70%.

Table 21: Prediction accuracy operational variables

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | |
|---|---|---|
| **Lag 1 Operational variables (75%)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | 2110 (73.04%) | 867 (23.47%) |
| **Non-Financial Distress** | 779 (26.96%) | 2827 (76.53%) |
| **Lag 2 Operational variables (68.19%)** | | |
| **Financial Distress** | 1628 (65.75%) | 978 (29.95%) |
| **Non-Financial Distress** | 848 (34.25%) | 2287 (70.05%) |
| **Lag 3 Operational variables (63.79%)** | | |
| **Financial Distress** | 1211 (58.45%) | 915 (32.30%) |
| **Non-Financial Distress** | 861 (41.55%) | 1918 (67.70%) |

### 5.2.5. Combined variables full sample

The final logistic regression for the full sample is the combined variables regression. The coefficients are noted in Table 22. As mentioned in the methodology the Altman, financial and operational variables are combined into one model. The operating results over total assets (ORTA), book value of equity over total assets (BVETA), current ratio (CR) and result ratio (RR) are removed from the regression because they had a high correlation with working capital (WCTA) and EBIT over total assets (EBITTA).

WCTA, LnRevenue and the Financial Distress lag are the perfect example for good predicting variables. The variables are significant at the 1%-level for all lags, have the right sign, and the strength of the coefficient is increasing when the lag decreases. If WCTA and LnRevenue increase, the probability of FD decreases. A Financial Distress lag dummy of 1, one period before t = 0 increases the probability of FD.

The sign of the coefficients for RTA, Competition and BVPV should be negative. If these variables increase, the chance of an HCI going into Financial Distress should decrease and not increase as the sign suggests.

EBITTA and RTA are significant for all lags but do not have a clear trend over time. For EBITTA the strength of the coefficient decreased from lag 3 (-1.291) to lag 2 (-1.051) and increased from lag 2 (-1.051) to lag 1 (-1.465). And for the RTA the coefficients decreased instead of increased.

The Efficiency ratio, RevenueD, Staff turnover and annual statement are not significant for any lag

Table 22: Coefficients of the combined model

| *Altman, Financial & Operational variables* | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Working capital/Total Assets* | -0.663 | 0.000*** | -0.590 | 0.000*** | -0.419 | 0.008*** |
| *EBIT/Total Assets* | -1.465 | 0.000*** | -1.051 | 0.006*** | -1.291 | 0.001*** |
| *Revenue/ Total Assets* | 0.102 | 0.003*** | 0.113 | 0.001*** | 0.167 | 0.000*** |
| *Efficiency ratio* | -0.468 | 0.300 | -0.109 | 0.809 | 0.088 | 0.858 |
| *Debt/EBITDA* | 0.009 | 0.004*** | 0.006 | 0.050* | 0.008 | 0.019** |
| *LnRevenue* | -0.121 | 0.000*** | -0.129 | 0.000*** | -0.111 | 0.000*** |
| *RevenueD (%)* | -0.025 | 0.675 | -0.024 | 0.646 | -0.024 | 0.596 |
| *Absenteeism* | 8.211 | 0.000*** | 3.720 | 0.062* | 4.898 | 0.021** |
| *Staff turnover* | 0.321 | 0.172 | 0.346 | 0.129 | -0.032 | 0.894 |
| *Competition (x1,000,000)* | 3.96 | 0.006*** | 2.05 | 0.170 | 0.402 | 0.801 |
| *Book value/ Purchase value* | 0.123 | 0.516 | 0.150 | 0.443 | 0.445 | 0.026** |
| *Accounts receivable* | 2.071 | 0.003*** | 1.497 | 0.037 | 1.078 | 0.169 |
| *Annual statement* | 0.109 | 0.122 | -0.001 | 0.988 | -0.025 | 0.742 |
| *Financial Distress* | 1.860 | 0.000*** | 1.256 | 0.000*** | 0.869 | 0.000*** |
| *Constant* | 0.943 | 0.188 | 0.891 | 0.109 | -0.618 | 0.299 |

***= significant at the 1%-level, ** = significant at the 5%-level , *= significant at the 10%-level

Table 23 shows the prediction accuracy of the combined model. This prediction accuracy is approximately the same as the financial and the operational models. The percentage of correctly classified HCI are respectively, 74.67%, 68.10% and 63.54% for one, two and three years before the event date. This model is able to correctly classify 74.67% of the HCI in the dataset.

This is good prediction accuracy, because it's significantly higher than 50%, and could be used for future predictions. The prediction accuracy for institutions in FD for the one, two and three-year lags are 71.42%, 63.42% and 49.17% and for institutions not in FD these are 77.13%, 71.52% and 74.22%.  The predictive power for predicting HCI in FD is not good for combined variables three years before the event date (49.17%). I would not recommend using the model in this situation. Finally, there is not a clear trend for the prediction accuracy of HCI, not in FD. The accuracy decreases from 74.22% to 71.52% and then increases to 77.13%.

Table 23: Prediction accuracy combined model

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | |
|---|---|---|
| **Lag 1 combined variables (74.67%)** | **Financial Distress** | **Non-Financial Distress** |
| **Financial Distress** | 1767 (71.42%) | 748 (22.87%) |
| **Non-Financial Distress** | 707 (28.58%) | 2523 (77.13%) |
| **Lag 2 Combined variables (68.10%)** | | |
| **Financial Distress** | 1316 (63.42%) | 807 (28.48%) |
| **Non-Financial Distress** | 759 (36.58%) | 2027 (71.52%) |
| **Lag 3 Combined variables (63.54%)** | | |
| **Financial Distress** | 860 (49.17%) | 606 (25.78%) |
| **Non-Financial Distress** | 889 (50.83%) | 1745 (74.22%) |

## 5.3. Results Altman model and logistic regression sub samples

The robustness check is done by dividing the sample in two equal periods. The first period is from 2009-2013 (I) and the second period is from 2014-2018 (II). The goal of the robustness check is to see if the same results are found for different periods, or that the results are period dependent. For the robustness test the Altman model and the combined logistic model will be tested. To make the comparison easier, the same variables as for the full sample are used. So, only the variables that were not highly correlated with other variables. The results of period I have a white background and the results of period II have a grey background.

### *5.3.1. Altman model subsample*

In Table 24 the results of the Altman model for period I and period II are noted. The prediction accuracy of the Altman model for period I is better than the full sample, while period II is worse than the full sample. The correctly classified institutions for the one, two and three-year lag are respectively 68.87%, 67.22% and 67.99% for period I and 60.58%, 58,46% and 57.84% for period II. In period II it’s harder to predict financial distress than in period I, but the results are still a lot better than 50%. The highest prediction accuracy for correctly classifying firms in financial distress is 83.10% for the period I sample. Although the hit ratio is higher than 50% for period I it is still not close to the prediction rate of previous papers. There is no clear trend in the prediction accuracy for institutions not in FD. The prediction accuracy decreased from lag 3 (66.52%) to lag 2 (62.89%) and then increased to lag 1 (63.82%). The rest of the results are in line with the expectations.

Table 24: Results Altman model 2009-2013 and 2014-2018

The left side of the column [a] is for period I and the right-side [b] is for period II.

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 Altman model (68.87%[a] / 60.58%[b])** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 300 (83.10%)[a] | 196 (76.18%)[b] | 368 (36.18%)[a] | 513 (42.89%)[b] |
| **Non-Financial Distress** | 61 (16.90%) | 59 (23.14%) | 649 (63.82%) | 683 (57.11%) |
| **Lag 2 Altman model (67.22% / 58.46%)** | | | | |
| **Financial Distress** | 222 (78,45%) | 203 (72.24%) | 272 (37,11%) | 531 (44.81%) |
| **Non-Financial Distress** | 61 (21,55%) | 78 (27.76%) | 461 (62,89%) | 654 (55.19%) |
| **Lag 3 Altman model (67.99% / 57.84%)** | | | | |
| **Financial Distress** | 134 (71,66%) | 216 (71.05%) | 157 (33,48%) | 536 (45.58%) |
| **Non-Financial Distress** | 53 (28,34%) | 88 (28.95%) | 312 (66,52%) | 640 (54.42%) |

5.3.2. Altman variables logistic regression subsample

Table 25 shows the coefficients of the Altman variables for period I and period II. The WCTA coefficient is the strongest in period I. Period II, lag 3 is the first time the WCTA coefficient is not significant at any level. The WCTA has a clear increase in coefficient strength from lag 3 to lag 1, and the sign of the coefficient is in line with the theory. EBITTA is significant for all lags in both periods. The EBITTA coefficient is the highest in the second period. The sign is also in line with the theory. The RTA coefficient is not significant in the first period for any lag, but is significant at the 1%-level in the second period. Just like the full sample, the sign of the coefficient is not in line with the expectations.

Table 25: Coefficients Altman model 2009-2013 and 2014-2018

| ***Altman variables 2009-2013*** | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Working Capital/Total Assets* | -2.273 | 0.000*** | -1.738 | 0.000*** | -1.682 | 0.000*** |
| *EBIT/Total Assets* | -3.947 | 0.000*** | -2.000 | 0.000*** | -2.162 | 0.000*** |
| *Revenue/Total Assets* | 0.047 | 0.174 | 0.039 | 0.322 | 0.040 | 0.439 |
| *Constant* | 0.468 | 0.000*** | 0.138 | 0.183 | 0.032 | 0.809 |
| ***Altman variables 2014-2018*** | | | | | | |
| *Working Capital/Total Assets* | -1.414 | 0.000*** | -0.845 | 0.000*** | -0.296 | 0.203 |
| *EBIT/Total Assets* | -6.60 | 0.000*** | -3.979 | 0.000*** | -3.040 | 0.000*** |
| *Revenue/Total Assets* | 0.372 | 0.000*** | 0.332 | 0.000*** | 0.386 | 0.000*** |
| *Constant* | 0.396 | 0.000*** | -0.043 | 0.712 | -0.452 | 0.002*** |

***= significant at the 1%-level, ** = significant at the 5%-level , *= significant at the 10%-level

Table 26 shows the prediction accuracy for the Altman model in both sub samples. Period I is better in predicting HCI in FD and less good in predicting institutions not in financial distress in comparison to the full sample. Period II and the full sample almost have the same structure, but Period II has a slightly better rate for correctly classifying HCI for one, two and three years before the event date. The rate for correctly classified institutions in Period I, for the one, two and three-year lag are 64.65%, 60.70% and 61.28%, and for period II these percentages are 65.30%, 61.66% and 61.70%. The sub samples have a low prediction accuracy for institutions in FD, but a high rate for HCI not in FD, just like the full sample. Also, the prediction accuracy increases, for correctly predicting institution not in FD, the further the variable is from the event date.

The subsamples give the same results as the full sample. The Altman model in the full sample is better in predicting institutions in FD, while the logistic regression in the full sample is better at predicting institutions not in FD.

Table 26: Prediction rates Altman variables for the logistic regression

The left side of the column a is for period I and the right-side b is for period II.

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 Altman model (64.65%a / 65.30%b)** | **Financial Distress** | | **Not Financial Distress** | |
| **Financial Distress** | 799 (56.71%)a | 489 (39.66%)b | 485 (28.72%)a | 232 (14.68%)b |
| **Not Financial Distress** | 610 (43.29%) | 744 (60.34%) | 1204 (71.28%) | 1348 (85.32%) |
| **Lag 2 Altman model (60.70% / 61.66%)** | | | | |
| **Financial Distress** | 389 (38.82%) | 244 (27.29%) | 285 (22.21%) | 144 (12.23%) |
| **Not Financial Distress** | 613 (61.18%) | 650 (72.71%) | 998 (77.79%) | 1033 (87.77%) |
| **Lag 3 Altman model (61.28% / 61.70%)** | | | | |
| **Financial Distress** | 155 (25.45%) | 104 (19.15%) | 121 (13.81%) | 75 (9.39%) |
| **Not Financial Distress** | 454 (74.55%) | 439 (80.85%) | 755 (86.19%) | 724 (90.61%) |

5.3.3. Combined variables subsample

The WCTA, LnRevenue, competition and the Financial Distress lag in the first period are the same as the full sample. They have the same sign and the same significance level in period I as in the full sample.

A lot of the variables that were significant in the full sample are not significant anymore in period I. DEBITDA is only significant for the first lag on the 10%-level while it was significant on all levels for the full sample. Although, RevenueD was not significant on any level in the full sample, it was only significant for the third lag in the first period. This would indicate that a change in revenue, three years before the date of the event has an effect on the status of the institution and one and two years before does not. Absenteeism is highly significant for the first lag just like the full sample, lost its significance for the second and third lag. In the full sample these two were still significant. Staff turnover is not significant in the full sample, but is significant for the second and third lag of period I. Just as the RevenueD, this finding is not in line with the expectations. Competition in the first period is significant for all lags at the 1%-level and increases when the lag decreases. Competition of the full sample was only significant at the first lag, with a significance level at 1%. Accounts receivable in the first period is significant at the 10%-level and 5%-level for the one and two-year lag. For the full sample this variable was only significant at the 1%-level for the first lag. Finally, the annual statement is not significant in the full sample, but significant at the 10%-level three years before the event date in period I. The ER and the BVPV are not significant in the first period. None of the signs are different in the first period compared to the full sample.

In table 28 the regression coefficients for the combined variables in period II can be found. WCTA loses its significance for the first two lags, but is significant at the 10%-level three years before the event date. EBITTA has a highly significant coefficient for the
first two lags but loses it's significant in the third lag. In the full sample, EBITTA was significant for all the lags. The efficiency ratio in the full sample did not have any effect on the regression outputs, but for period II this variables is significant at the 5%-level for lag 1 and 2. It's noteworty that the DEBITDA ratio is significant for all three lags in the full sample, but only significant at the 10%-level period II. LnRevenue loses its strength in the second period. In the full sample, this variable was significant for all lags, but in the second period it is only significant at the 5%-level and only for the first lag. Just like in the first period Absenteeism is only significant for the first lag, while it is significant for all lags in the full sample. Staff turnover has a small significance in the first and second period, but was not significant for the full sample. Competition is really significant in the first period, but loses it's significance in the second period, the first and third lag have significant coefficients at the 10% and the 5%-level. The BVPV has a small significance level for period II, and is significant for the third lag at the 5%-level in the full sample.

Table 27: Results combined variables logistic regression 2009-2013

| ***Altman, Financial & Operational variables 2009-2013*** | *Coefficient (Lag 1)* | *P-value* | *Coefficient (Lag 2)* | *P-value* | *Coefficient (Lag 3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Working capital/Total Assets* | -1.168 | 0.000*** | -1.137 | 0.000*** | -1.289 | 0.001*** |
| *EBIT/Total Assets* | -0.836 | 0.150 | -0.037 | 0.957 | -1.490 | 0.108 |
| *Revenue/ Total Assets* | 0.175 | 0.756 | -0.007 | 0.911 | 0.103 | 0.266 |
| *Efficiency ratio* | -1.089 | 0.163 | -0.409 | 0.654 | -0.913 | 0.466 |
| *Debt/EBITDA* | 0.009 | 0.079* | 0.005 | 0.466 | 0.004 | 0.647 |
| *LnRevenue* | -0.138 | 0.006*** | -0.230 | 0.000*** | -0.164 | 0.038** |
| *RevenueD* | 0.384 | 0.118 | 0.507 | 0.114 | 0.893 | 0.036** |
| *Absenteeism* | 9.114 | 0.004*** | 3.403 | 0.368 | 7.826 | 0.113 |
| *Staff turnover* | 0.439 | 0.108 | 0.654 | 0.030** | 0.708 | 0.088* |
| *Competition (x10,000)* | 0.113 | 0.000*** | 0.111 | 0.000*** | 0.074 | 0.043** |
| *Book value/ Purchase value* | -0.259 | 0.930 | -0.329 | 0.305 | -0.239 | 0.544 |
| *Accounts receivable* | 2.180 | 0.075* | 2.724 | 0.058** | 3.171 | 0.103 |
| *Annual statement* | 0.192 | 0.140 | -0.000 | 0.999 | -0.415 | 0.075* |
| *Financial Distress lag* | 2.165 | 0.000*** | 1.685 | 0.000*** | 1.085 | 0.000*** |
| *Constant* | 1.151 | 0.000*** | 2.019 | 0.073* | 2.252 | 0.133 |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

The coefficients for Accounts receivable in the second period are not significant at all. In the full sample, the variable is only significant for the first lag, at the 1%- level. The annual statement is not significant in the full sample, but is significant at the 1%-level for the first lag of period II.

It's not quite clear how variables that are not significant in both subsamples, end up being significant in the full sample. RTA and the Financial Distress lag are the only variables that are significant for all lags at the 1%-level. RevenueD and Accounts receivable are not significant on any level. None of the variables changed sign in the second period.

Table 28 shows the prediction accuracy for the combined variables in period I and period II. Period I has an overall better classification rate than period II. The highest correctly classified rate is 77.61% for predicting the status of an HCI in period I. Period I is better in predicting firms in financial distress, and period II is on average better in predicting firms not in financial distress. Also, it is noteworthy that the prediction accuracy of institutions in period II, that are not in FD increases, when the lag increases. All the other percentages are in line with the expectations. Although the coefficients of the first and second period have the same sign as the full sample, most of the coefficients are a lot less significant. The robustness check does not hold.

Table 28: Results combined variables logistic regression 2014-2018

| ***Altman, Financial & Operational variables*** | *Coefficient (L1)* | *P-value* | *Coefficient (L2)* | *P-value* | *Coefficient (L3)* | *P-value* |
|---|---|---|---|---|---|---|
| *Working capital/Total Assets* | -0.362 | 0.166 | -0.111 | 0.725 | 0.783 | 0.073* |
| *EBIT/Total Assets* | -2.666 | 0.000*** | -2.378 | 0.006*** | 0.728 | 0.530 |
| *Revenue/ Total Assets* | 0.289 | 0.000*** | 0.390 | 0.000*** | 0.324 | 0.005*** |
| *Efficiency ratio* | -1.568 | 0.042** | -2.311 | 0.011** | -1.319 | 0.308 |
| *Debt/EBITDA* | 0.008 | 0.083* | 0.003 | 0.548 | 0.002 | 0.793 |
| *LnRevenue* | -0.111 | 0.015** | -0.078 | 0.150 | -0.109 | 0.165 |
| *RevenueD* | -0.572 | 0.166 | -0.206 | 0.587 | -0.274 | 0.594 |
| *Absenteeism* | 9.128 | 0.010** | 6.215 | 0.150 | 1.304 | 0.828 |
| *Staff turnover* | 0.458 | 0.408 | 1.041 | 0.092* | 0.838 | 0.327 |
| *Competition (x10,000)* | -0.052 | 0.062* | -0.052 | 0.157 | -0.132 | 0.031** |
| *Book value/ Purchase value* | 0.687 | 0.059* | 0.753 | 0.087* | 1.373 | 0.030** |
| *Accounts receivable* | 0.409 | 0.725 | -0.595 | 0.661 | 0.040 | 0.984 |
| *Annual statement* | 0.381 | 0.017*** | 0.091 | 0.620 | 0.240 | 0.318 |
| *Financial Distress lag* | 1.583 | 0.000*** | 1.073 | 0.000*** | 1.167 | 0.000*** |
| *Constant* | 1.131 | 0.246 | 1.083 | 0.351 | 0.262 | 0.875 |

***= significant at the 1%-level, ** = significant at the 5%-level, *= significant at the 10%-level

Table 29: Results combined variables 2009-2013 and 2014-2018

The left side of the column a is for period I and the right-side b is for period II.

| **Actual Group Membership (0.5)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 combined variables (77.61%a / 72.49%b)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 746 (76.99%)a | 593 (66.63%)b | 272 (21.90%)a | 271 (23.06%)b |
| **Non-Financial Distress** | 223 (23.01%) | 297 (33.37%) | 970 (78.10%) | 904 (76.94%) |
| **Lag 2 combined variables (72.79% / 67.89%)** | | | | |
| **Financial Distress** | 405 (69.47%) | 288 (53.14%) | 207 (24.88%) | 176 (22.08%) |
| **Non-Financial Distress** | 178 (30.53%) | 254 (46.86%) | 625 (75.12%) | 621 (77.92%) |
| **Lag 3 combined variables (66.01% / 67.55%)** | | | | |
| **Financial Distress** | 174 (57.81%) | 120 (47.81%) | 109 (25.65%) | 78 (19.85%) |
| **Non-Financial Distress** | 127 (42.19%) | 131 (52.19%) | 316 (74.35%) | 315 (80.15%) |

# 6. Conclusion & Further research

This paper tries to classify institutions into a prespecified group with the help of the Altman, financial and operational variables. The results are not as expected. Earlier papers had much higher prediction rates with the Altman test and the logistic regressions. An explanation for the lower prediction rate is, that this paper tries to forecast financial distress instead of bankruptcy. Bankruptcy has a more certain outcome and is definite. With the right intervention, a financial distresses institution can be turned around. A reason for the low prediction accuracy of the Altman variables could be, that the operating results over total assets (ORTA) and revenue over total assets (RTA) were used as proxies in the revised Altman model.

There was a clear difference between most of the variable means of HCI in FD, and institutions that were not. Looking at the significance of the t-test, one would think that the models would have good prediction powers. Altman (1983) said that prediction accuracies can differ over different institutions. Almwajeh (2004) found in his paper, that both the discriminate analyses and the logistic regression models were able to predict the financial status of the hospitals, with a prediction power of respectively 90.2% and 100%. This is the only paper I could find that used the Altman model and logistic regression on hospitals, so it's difficult to compare the results to earlier papers.

The variables that had the best significant levels for the one, two and three year lag , were WCTA, EBITTA, RTA, LnRevenue and Financial Distress lag. It is important for health care institutions to focus on these variables to keep their institution from going into FD. The best prediction model for the full sample was formed by the operational variables. The combined model for period II had the highest percentage of correctly classified institutions overall. The percentage for correctly classifying HCI for the operational variables in the full sample, and the combined variable in period II are respectively, 75% and 77.61%. Although earlier papers had higher prediction rates, I would still recommend using the models in this thesis, with a prediction accuracy of more than 70%. It is not the best prediction accuracy, but it is better than simply guessing the outcome.

The significance of the variables over the two periods were completely different from one another. The Altman variables were stronger in the first period and the variables from Finance Ideas were stronger in the second period. Period I had better prediction accuracies for the Altman model, and had approximately the same percentage of correctly classified HCI, for the logistic regression of the Altman variables as period II. The rate for correctly classified institutions was higher for period I than for period II in the first and second lag, but almost the same for the third lag.

A limitation of this research is that there are not many studies about bankruptcy/financial distress for HCI. Therefor it is hard to compare the results. Using proxies for the Altman variables, could bias the outcome of the Altman test and give a lower prediction accuracy. To the extent of my knowledge, there was not a single paper about FD that used the variables I used for this research, which made it also difficult to compare the outcomes. It is unclear if the results differ, because the FD for this industry is harder to predict, or that the models in combination with the variables used were just not good enough. The papers on bankruptcy are written for different industries, different time periods and different variables, so, it is difficult to compare the findings of the thesis with earlier research. Another limitation was that Altman did not explain the way he created his model, so I could not reproduce it with my own variables.

The last limitation for me was to decide what to add to the thesis, and what I should simply leave out.

For future research I would recommend the use of other variables. The variables used for this research were coming from Finance Ideas. The only variable from FI that was used in earlier research was the Current ratio (CR). So, again this made it difficult to compare the results of the paper to previous research. It is also interesting to look at other countries and try to use other models like the Hazard model or Neural Network Model. Instead of accounting-based, financial and operational variables, other variables can be used for future research. Some examples are macroeconomic variables or corporate governance indicators. With the findings and limitations of this thesis, I encourage further research in this underexposed area.

# References


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

Almwajeh, O. (2004). Applying Altman's Z-Score model of bankruptcy for the prediction of financial distress of rural hospitals in Western Pennsylvania (Doctoral dissertation, Indiana University of Pennsylvania).

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

Altman, E. I. (1983). Multidimensional graphics and bankruptcy prediction: a comment. *Journal of Accounting Research*, 297-299.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2014). Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Available at SSRN 2536340*.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting*, *28*(2), 131-171.

Altman, E. I., & Loris, B. (1976). A financial early warning system for over-the-counter broker-dealers. *The Journal of Finance, 31*(4), 1201-1217.

Altman, E. I., & McGough, T. P. (1974). Evaluation of a company as a going concern. *Journal of Accountancy*, *138*(6), 50-57

Anjum, S. (2012). Business bankruptcy prediction models: A significant study of the Altman’s Z-score model. *Available at SSRN 2128475*.

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

Beaver, W. H. (1968). The information content of annual earnings announcements. *Journal of accounting research*, 67-92.

Blum, M. P. (1969). *The failing company doctrine* (Doctoral dissertation, UMI Dissertation Services).

Collins, R. A., & Green, R. D. (1982). Statistical methods for bankruptcy forecasting. *Journal of Economics and Business*, *34*(4), 349-354.

Dambolena, I. G., & Khoury, S. J. (1980). Ratio stability and corporate failure. *The Journal of Finance*, *35*(4), 1017-1026.

Darayseh, M., Waples, E., & Tsoukalas, D. (2003). Corporate failure for manufacturing industries using firms specifics and economic environment with logit analysis. *Managerial Finance*.

Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. Journal of accounting research, 167-179.

Edmister, R. O. (1972). An empirical test of financial ratio analysis for small business failure prediction. Journal of Financial and Quantitative analysis, 7(2), 1477-1493.

Gu, Z., & Gao, L. (2000). A multivariate model for predicting business failures of hospitality firms. *Tourism and Hospitality Research*, *2*(1), 37-49.

Hand, D. J. (1981). Discrimination and Classification. New York: John Wiley and Sons

Huberty, C. J. (1984). Issues in the use and interpretation of discriminant analysis. *Psychological Bulletin*, *95*(1), 156.

Kim, H., & Gu, Z. (2006). A logistic regression analysis for predicting bankruptcy in the hospitality industry. *The Journal of Hospitality Financial Management*, *14*(1), 17-34.

Lacher, R. C., Coats, P. K., Sharma, S. C., & Fant, L. F. (1995). A neural network for classifying the financial health of a firm. *European Journal of Operational Research*, *85*(1), 53-65.

Laitinen, E. K., & Laitinen, T. (2000). Bankruptcy prediction: Application of the Taylor's expansion in logistic regression. *International review of financial analysis*, *9*(4), 327-349.

Lau (1987) Lau, A.H. (1987), A five state financial distress prediction model. Journal of Accounting Research, 25, pp 127-138.

Merwin, C. L. (1942). *Financing small corporations in five manufacturing industries, 1926-1936*. National Bureau of Economic Research, New York.

Morey, J., Scherzer, G., & Varshney, S. (2004). Predicting financial distress and bankruptcy for hospitals. Journal of Business & Economics Research (JBER), 2(9).

Moyer, R. C. (1977). Forecasting financial failure: a re-examination. *Financial Management (pre-1986)*, *6*(1), 11.

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

Roy A. Foulke, Practical Financial Statement Analysis, 5th Ed., (New York, McGraw-Hill, 1961)

Russ, R. W., Peffley, W. W., & Greenfield, A. C. (2004). The Altman Z-score revisited. *American Accounting Association*.

Zavgren, C. V., & Friedman, G. E. (1988). Are bankruptcy prediction models worthwhile? An application in securities analysis. *Management International Review*, 34-44.

Appendix

Table 1: Summary of previous papers

| Paper | Method | Industry/number of firms | Important ratios | Prediction accuracy | Time period |
|---|---|---|---|---|---|
| Beaver (1966) | Univariate model | Big asset size firms | Annual cashflow/Total debt | 13% and 22% one year and 5 years before bankruptcy | 1954-1964 |
| Blum (1969) | | 230 industrial firms | | One, two and five years prior to bankruptcy is 93-95%, 80% and 70% | 1954-1968 |
| Altman (1968) | MDA | 66 firms manufacturing industry. | Working capital/Total assets, retained earnings /Total assets, earnings before interest and taxes/Total Assets, market value of equity/book value of total debt, and sales/total assets | 79% one year before bankruptcy. | 1946-1965 |
| Deakin (1972) | Discriminant analyses and Dichotonomous classification test | 32 failed firms | 14 financial ratios | Error rates were respectively 22%, 6%, 12%, 23%, and 15% for one-five years prior to bankruptcy | 1964-1970 |
| Altman et al. (1976) | MDA | 40 bankrupt firms and 113 healthy firms | Net income/Total assets, Total liabilities+Subordinate loans/Owners equity, Total assets/adjusted net capital, (ending capital-capital addition)/Beginning Capital, Scaled age, and composite. | 90% for failed firms and 90.3% for healthy firms | 1970-1972 |
| Gu and Gao (2000) | MDA | 28 firms, 14 healthy and 14 bankrupt. | Total liablities/Total assets, EBIT to current liabilities, gross profit margin, long-term debt to total assets, and sales to fixed assets | 93% one year before failure. | 1987-1996 |

| | | | | | |
|---|---|---|---|---|---|
| Russ et al. (2004) | | | | Type I and Type II error are 20.6% and 28.4% | |
| Dambolena and Khoury (1980) | MDA | 68 firms. 34 failed and 34 healthy firms. | Same as for the original Altman (1968) model. | Prediction accuracy for one, three and five years are respectively, 87%, 85% and 78% | 1969-1975 |
| Altman (1983) | Revised Z-model | Unlisted industrial (non-financial) companies, limited owner liability, big enough company size. | Working capital/Total assets, Retained earnings/Total assets, Earnings before interest and taxes/ Total assets, Book value of total liabilities, and Sales/Total assets | Able to correctly predict 90.1% of the bankrupt firms and 97% of the healthy firms. | |
| Altman (2014) | Revised Z-model | 34 countries, with 31 European countries and 3 non-European countries. Private and no-financial firms across all industries. | Working capital/Total assets, Retained earnings/Total assets, Earnings before interest and taxes/ Total assets, Book value of total liabilities, and Sales/Total assets | Predictive power was around 75% with some over 90%. | 2007-2010 |
| Ohlson (1980) | Logistic regression | 105 bankrupt firms and 2058 non-bankrupt firms. Must be classified as an industrial and traded over the counter. | Size(size (log(total assets/GNP price-level index), total liabilities divided by total assets, working capital divided by total assets, current liabilities divided by current assets, a dummy were one if total liabilities exceed total assets, zero otherwise, net income divided by total assets, funds provided by operations divided by total liabilities, a dummy were one if the net income was negative the last two years, zero otherwise, and a change in net income. | Two years before bankruptcy the prediction accuracy was 92% | 1970-176 |
| Zavrgen (1988) | Logistic regression | 45 bankrupt firms matched with healthy | Inventory Turnover, Receivables Turnover, Cash Position, Short-term liquidity, Return on | The minimal total classification rates for 1 to 5 years prior to | 1979-1980 |

| | | | | | |
|---|---|---|---|---|---|
| | | firms of the same industry and asset size | Investment, Financial Leverage, and Capital Turnover. | bankruptcy were 18%, 17%, 28%, 27% and 20%. | |
| Darayseh et al. (2003) | Logit analyses | 220 firms of which 110 firms were healthy and 110 were non-healthy firms. | Profit margin on sales, return on investment, times interest earned, debt/equity, quick ratio, accounts receivable ratio, gross national product, interest rate and stock price index. | Prediction power fort he in-sample and out of sample were respectively, 87.82% and 89.50% | 1990-1997 |
| Lau (1987) | Logit analyses | Each sample contained 350 firms | Loan restrictive terms, industry normalized debt/equity, working capital flow/Total debt, trend of common stock price, industry normalized operating expenses/sales, distribution of common stock dividends, liquidation of operating assets, trend of capital expenditure, trend of working capital flow, omission or reduction of dividend payments | The original sample had a prediction accuracy of 96%, 92%, and 90% one two and three years before bankruptcy. For the holdout sample the probabilities were 80%, 79% and 85% for one, two and three years. | 1971-1977 |
| Kim & Gu (2006) | Logit model | Hospitality industry. 32 firms for which 16 bankrupt firms and 16 healthy firms. | (Liquidity) Current ratio, quick ratio and operating cash flows to current liabilities. (Solvency) Debt ratio, long-term debt to total capitalization, operating cash flow/ total liabilities, times interest earned ratio (Profitability) gross profit margin, net profit margin, gross return on assets, and return on assets (Efficiency) total asset turnover and fixed asset turnover | Prediction accuracy were 91% and 84% for one and two years before the event took place. | 1999-2004 |
| Morey et al. (2004) | | Five upstate New York hospitals | Liquidity (Cash-flow to Total Debt), leverage long-term debt per bed, profitability (equity financing) and efficiency. | | 1998-2001 |
| Almwajeh (2004) | Altman Z-score model/ Logistic regression | 65 rural hospitals | Liquidity, profitability, efficiency and productivity | Prediction accuracy of 90.2% fort the MDA and 100% for the logistic regression. | |

Table 17A: Altman logistic regression results for 0.45 and 0.55 cut-off point

The left side of the column a is for cut-off point 0.45 and the right-side b is cut-off point 0.55.

| **Actual Group Membership (0.45a/ 0.55b)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 Altman model (63.83%a / 66.64%b)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 1819 (62.06%a) | 846 (28.86%b) | 1301 (34.79%a) | 407 (10.88%b) |
| **Non-Financial Distress** | 1112 (37.94%) | 2085 (71.14%) | 2439 (65.21%) | 3333 (89.12%) |
| **Lag 2 Altman model (60.79% / 60.31%)** | | | | |
| **Financial Distress** | 1295 (51.55%) | 399 (15.88%) | 1067 (32.21%) | 199 (6.01%) |
| **Non-Financial Distress** | 1217 (48.45%) | 2113 (84.12%) | 2246 (67.79%) | 3114 (93.99%) |
| **Lag 3 Altman model (60.43% / 59.85%)** | | | | |
| **Financial Distress** | 858 (40.76%) | 233 (11.07%) | 726 (25.20%) | 130 (4.51%) |
| **Non-Financial Distress** | 1247 (59.24%) | 1872 (88.93%) | 2155 (74.80%) | 2751 (95.49%) |

Table 19A: Results Financial variables for 0.45 and 0.55 cut-off point

The left side of the column a is for cut-off point 0.45 and the right-side b is cut-off point 0.55.

| **Actual Group Membership (0.45a / 0.55b)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 financial variables (74.61%a / 74.66%b)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 1790 (71.34%a) | 1783 (71.06%b) | 759 (22.92%a) | 749 (22.61%b) |
| **Non-Financial Distress** | 719 (28.66%) | 726 (28.94%) | 2553 (77.08%) | 2563 (77.39%) |
| **Lag 2 financial variables (68.15% / 67.12%)** | | | | |
| **Financial Distress** | 1347 (64.02%) | 1167 (55.47%) | 830 (28.84%) | 701 (24.36%) |
| **Non-Financial Distress** | 757 (35.98%) | 937 (44.53%) | 2048 (71.16%) | 2177 (75.64%) |
| **Lag 3 financial variables (63.84 % / 61.83%)** | | | | |
| **Financial Distress** | 1020 (57.46%) | 649 (36.56%) | 753 (31.44%) | 389 (16.24%) |
| **Non-Financial Distress** | 755 (42.54%) | 1126 (63.44%) | 1642 (68.56%) | 2006 (83.76%) |

Table 21A: Results Operational variables for 0.45 and 0.55 cut-off point

The left side of the column $_a$ is for cut-off point 0.45 and the right-side $_b$ is cut-off point 0.55.

| **Actual Group Membership (0.45$_a$ / 0.55$_b$)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 operational variables (75.00%$_a$ / 75.00%$_b$)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 2110 (73.04%$_a$) | 2110 (73.04%$_b$) | 867 (23.47%$_a$) | 867 (23.47%$_b$) |
| **Non-Financial Distress** | 779 (26.96%) | 779 (26.96%) | 2827 (76.53%) | 2827 (76.53%) |
| **Lag 2 operational variables (68.19% / 68.19%)** | | | | |
| **Financial Distress** | 1628 (65.75%) | 1624 (65.69%) | 978 (29.95%) | 974 (29.83%) |
| **Non-Financial Distress** | 848 (34.25%) | 852 (34.41%) | 2287 (70.05%) | 2291 (70.17%) |
| **Lag 3 operational variables (63.91% / 62.47%)** | | | | |
| **Financial Distress** | 1244 (60.04%) | 948 (45.75%) | 942 (33.25%) | 717 (25.31%) |
| **Non-Financial Distress** | 828 (39.96%) | 1124 (54.25%) | 1891 (66.75%) | 2116 (74.69%) |

Table 23A: Results Combined variables for 0.45 and 0.55 cut-off point

The left side of the column $_a$ is for cut-off point 0.45 and the right-side $_b$ is cut-off point 0.55.

| **Actual Group Membership (0.45$_a$ / 0.55$_b$)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 combined variables (74.62% / 74.60%)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 1768 (71.46%$_a$) | 1757 (71.02%$_b$) | 752 (22.99%$_a$) | 742 (22.68%$_b$) |
| **Non-Financial Distress** | 706 (28.54%) | 717 (28.98%) | 2519 (77.01%) | 2529 (77.32%) |
| **Lag 2 Combined variables (68.00% / 67.71%)** | | | | |
| **Financial Distress** | 1332 (64.19%) | 1158 (55.81%) | 828 (29.22%) | 668 (23.57%) |
| **Non-Financial Distress** | 743 (35.81%) | 917 (44.19%) | 2006 (70.78%) | 2166 (76.43%) |
| **Lag 3 Combined variables (64.02% / 63.17%)** | | | | |
| **Financial Distress** | 1001 (57.23%) | 611 (34.39%) | 727 (30.92%) | 372 (15.82%) |
| **Non-Financial Distress** | 748 (42.77%) | 1138 (65.07%) | 1624 (69.08%) | 1979 (84.18%) |

Table 26A: Results Altman variables logistic regression 2009-2013

The left side of the column $_a$ is for cut-off point 0.45 and the right-side $_b$ is cut-off point 0.55.

| **Actual Group Membership (0.45$_a$ / 0.55$_b$)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 Altman model (63.11%$_a$ / 63.88%$_b$)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 988 (70.12%$_a$) | 575 (40.81%$_b$) | 722 (42.75%$_a$) | 285 (16.87%$_b$) |
| **Non-Financial Distress** | 421 (29.88%) | 834 (59.19%) | 967 (57.25%) | 1404 (83.13%) |
| **Lag 2 Altman model (60.18% / 59.87%)** | | | | |
| **Financial Distress** | 617 (61.58%) | 196 (19.56%) | 525 (40.92%) | 111 (8.65%) |
| **Non-Financial Distress** | 385 (38.42%) | 806 (80.44%) | 758 (59.08%) | 1172 (91.35%) |
| **Lag 3 Altman model (60.54% / 61.62%)** | | | | |
| **Financial Distress** | 292 (47.95%) | 84 (13.79%) | 269 (30.71%) | 45 (5.14%) |
| **Non-Financial Distress** | 317 (52.05%) | 525 (86.21%) | 607 (69.29%) | 831 (94.86%) |

Table 26B: Results Altman variables logistic regression 2014-2018

The left side of the column $_a$ is for cut-off point 0.45 and the right-side $_b$ is cut-off point 0.55.

| **Actual Group Membership (0.45$_a$ / 0.55$_b$)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **Lag 1 Altman model (66.33%$_a$ / 63.85%$_b$)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 690 (55.96%$_a$) | 366 (29.68%$_b$) | 404 (25.57%$_a$) | 150 (9.49%$_b$) |
| **Non-Financial Distress** | 543 (44.04%) | 867 (70.32%) | 1176 (74.43%) | 1430 (90.51%) |
| **Lag 2 Altman model (60.99% / 61.03%)** | | | | |
| **Financial Distress** | 371 (41.50%) | 166 (18.57%) | 285 (24.21%) | 79 (6.71%) |
| **Non-Financial Distress** | 523 (58.50%) | 728 (81.43%) | 892 (75.79%) | 1098 (93.29%) |
| **Lag 3 Altman model (63.49% / 61.33%)** | | | | |
| **Financial Distress** | 168 (30.94%) | 73 (13.44%) | 115 (14.39%) | 49 (6.13%) |
| **Non-Financial Distress** | 375 (69.06%) | 470 (86.56%) | 684 (85.61%) | 750 (93.87%) |

Table 29A: Results Combined variables 2009-2013

The left side of the column a is for cut-off point 0.45 and the right-side b is cut-off point 0.55.

| **Actual Group Membership (0.45a / 0.55b)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **L1 combined variables (77.61%a / 77.48%b)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 750 (77.40%a) | 739 (76.26%b) | 276 (22.22%a) | 268 (21.58%b) |
| **Non-Financial Distress** | 219 (22.60%) | 230 (23.74%) | 966 (77.78%) | 974 (78.42%) |
| **L2 combined variables (72.44% / 72.86%)** | | | | |
| **Financial Distress** | 433 (74.27%) | 364 (62.44%) | 240 (28.85%) | 165 (19.83%) |
| **Non-Financial Distress** | 150 (25.73%) | 219 (37.56%) | 592 (71.15%) | 667 (80.17%) |
| **L3 combined variables (66.94% / 66.53%)** | | | | |
| **Financial Distress** | 195 (64.78%) | 135 (44.85%) | 134 (31.53%) | 77 (18.12%) |
| **Non-Financial Distress** | 106 (35.22%) | 166 (55.15%) | 291 (68.47%) | 348 (81.88%) |

Table 29B: Results combined variables 2014-2018

The left side of the column a is for cut-off point 0.45 and the right-side b is cut-off point 0.55.

| **Actual Group Membership (0.45a / 0.55b)** | **Predicted Group Membership** | | | |
|---|---|---|---|---|
| **L1 combined variables (72.93%a / 72.64%b)** | **Financial Distress** | | **Non-Financial Distress** | |
| **Financial Distress** | 615 (69.10%a) | 580 (65.17%b) | 284 (24.17%a) | 255 (21.70%b) |
| **Non-Financial Distress** | 275 (30.90%) | 310 (34.83%) | 891 (75.83%) | 920 (78.30%) |
| **L2 Combined variables (66.02% / 68.71%)** | | | | |
| **Financial Distress** | 328 (60.52%) | 231 (42.62%) | 241 (30.24%) | 108 (13.55%) |
| **Non-Financial Distress** | 214 (39.48%) | 311 (57.38%) | 556 (69.76%) | 689 (86.45%) |
| **L3 Combined variables (65.53% / 63.56%)** | | | | |
| **Financial Distress** | 141 (56.18%) | 86 (34.26%) | 112 (28.50%) | 44 (11.20%) |
| **Non-Financial Distress** | 110 (43.82%) | 165 (65.74%) | 281 (71.50%) | 349 (88.80%) |