Norwegian School of Economics

Bergen, Fall 2017

Predicting bankruptcy for Norwegian firms

A study of Altman's Z´´-model using alternative ratios

Ibrahim Pelja & Tormod B. Stemland

Supervisor: Finn Kinserdal

Master thesis, MSc, Business Analysis and Performance Management

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

In this thesis, we study whether modern accounting ratios based on deductive reasoning and modern financial statements are superior to older, conventional ratios. The focus of this study is to evaluate to what extent alternative ratios can improve bankruptcy prediction models. This is done using Altman's revised Z´´-model as a base throughout the study.

To the best of our knowledge, this is the first study using this approach. We have found no studies that aim to improve the Z´´-model by replacing the ratios with alternative ratios that consider a similar aspect. Additionally, we found no studies that directly criticize the ratios applied by Altman. We find a general limitation on the subject of bankruptcy prediction to be a lack of reasoning behind the applied ratios.

We develop alternative models to the Z´´-model. These models are based on the outline of the Z´´-model and produced using the same statistical approach, namely multivariate discriminant analysis. Our models were developed using a sample of 158 Norwegian firms from 2009-2016. The sample consists of 79 bankrupt firms and 79 non-bankrupt firms.

In general, we find that a majority of the alternative ratios applied in the analysis improved the Z´´-model on an individual basis. We also highlight three alternative models that produce results superior to those of the Z´´-model. These models all consist of two alternative ratios and two of Altman's original ratios. Generally, we found Financial Assets/Liabilities to be a particularly good ratio. On the other hand, we found Working Capital/Total Assets, which was part of the original model, to be a poor ratio.

The findings of this study support our hypothesis that some modern ratios are better suited to predicting bankruptcy than conventional ratios.

Keywords: Bankruptcy prediction, multivariate discriminant analysis, Altman's Z´´-score, alternative accounting ratios.

Acknowledgements

This thesis was written in the fall of 2017 at Norwegian School of Economics, as part of our Master of Science degree in Economics and Business Administration.

The work on this thesis has been challenging, yet rewarding and informative. We hope our work contributes to the field of bankruptcy prediction, especially by putting emphasis on alternative ratios in prediction models.

We want to thank the Centre for Applied Research at NHH for providing us with the database that built the foundation of this research.

Most of all, we would like to express our gratitude to our supervisor, Associate Professor Finn Kinserdal. His support and guidance has been of great value, and has lead us to improve the quality of this research.

Bergen, December 2017

Ibrahim Pelja

Tormod B. Stemland

Table of contents

1. Introduction

Ratio analysis can be traced back to around year 300 B.C. and Euclid's analysis of ratio properties (Horrigan, 1968). The application of ratios has developed over time since then, and the use of ratios in finance is a recent phenomenon. As the industrial revolution changed the economic outlook and as professional managers emerged, the need for financial statement analysis was evident. Financial ratios soon became an important tool for assessing firms. Numerous financial ratios have emerged since then, with varying degree of sophistication and theoretical fundament. During the development of financial ratios, there has generally been two different focuses: performance analysis and credit risk analysis (Horrigan, 1968). The latter has been the dominant force in the development and is the focus of this study.

The growth of the financial sector and the increased demand for loans motivated the development of financial ratios that focus on predicting default risk (Horrigan, 1968). This in turn inspired the development of statistical models that aimed to predict bankruptcies. Bankruptcies are defined in various ways for different models, but most researchers use the legal, regional definition. The models apply different types of statistical methods and use a variety of ratios.

Despite the development in methods used in bankruptcy prediction, there has been limited advances in the ratios that are applied. The Current Ratios were for instance developed in the early 1900s and are still used today (Horrigan, 1968). Following this, we have to consider whether the conventional standard ratios used today are the most appropriate.

1.1 Motivation and objective of the study

Our objective is to investigate whether more sophisticated ratios, based on deductive reasoning, provide more information than the standard ratios that are commonly used today. We argue that many common ratios lack deductive logic and are for that reason suboptimal for analytical purposes. Since bankruptcy prediction is concerned with estimating the risk of default, better ratios are of important value as they reduce the risk of losses.

Our main objective is twofold. We want to examine whether modern ratios based on deductive logic are better at predicting the risk of default. In order to do this, we want to develop a bankruptcy prediction model using the same outline as one of Altman's three models, namely the revised Z´´-model, and test if our alternative model is superior. By doing this, we hope to find evidence for our hypotheses:

H_1: Modern ratios based on deductive reasoning and more comprehensive financial statements are superior to older conventional ratios, for predicting bankruptcies.

H_2: The modern ratios are able to produce a bankruptcy prediction model that is superior to Altman's Z´´-score model, which in this study is representative of older models.

By closely replicating Altman's study we want to develop and compare models that, to the best of our ability, keep the endogenous factors constant. These factors are the categories of the ratios, the number of ratios, and the procedure. This will enable us to compare the models with a certain degree of confidence that the differences in results are solely based on the change in ratios.

The results are compared to Altman's Z´´-model and evaluated with respect to their prediction accuracy in different samples. If the alternative models are superior we can argue that our hypotheses are correct.

To the best of our knowledge, there have been no studies that aim to improve the Z´´-model by replacing the ratios with alternative ratios that consider a similar aspect. Additionally, we found no studies that directly criticize the ratios that Altman applied.

This study is not concerned with evaluating ratios for the purpose of performance analysis. Additionally, we are not interested in producing the best model, but rather test of we can improve an already existing model using alternative ratios. We also do not consider alternative methods. This is because we want to replicate Altman's approach in order to better compare the results.

1.2 Limitations

In this section, the limitations of the study is presented. These are caused by a lack of available data, the structure of the study, and personal discretion, for instance limiting the study to Norwegian companies.

The accounting data is gathered from annual reports which report the financial situation as of the 31.12 of that year. Ideally one would want data with a time frame as close as possible to the bankruptcy in order to capture any effects from changes in the firm's financial position. However, there are no available databases that contain accounting data based on a shorter time frame.

Another limitation to the data is that some bankrupt companies tend to not file their annual report if the bankruptcy occurs within the first half of the year. This results in a loss of information and reduces the predictive power of the model. Additionally, explanatory notes were difficult to obtain for some companies, especially bankrupt firms. The lack of notes means that some assumptions need to be made about the quality of the data.

Market data is also omitted as most companies are privately held. The inclusion of market data could have improved the model as it may contain valuable, forward-looking information. However, this issue is of lesser importance as the objective of the study is not to develop the best overall model, but investigate whether our alternative model is better than Altman's Z´´-model.

We also need to consider that we use Norwegian companies as opposed to American companies. Structural differences between the business environments between these countries may affect the suitability of different ratios.

1.3 Outline

The structure of the thesis is as follows. In the next chapter, we first review earlier research on the topic of bankruptcy prediction. This is followed by a thorough review of the research related to Altman's models.

Next, we describe the methodology used to develop and test the alternative models and their results. The models are derived using multivariate discriminant analysis (MDA).

Subsequently, Altman's Z``-model is presented in detail. The following section provides a detailed criticism of Altman's ratios and a presentation of our alternative ratios. This is the longest section by design because we want to emphasize the importance of using deductive reasoning when choosing ratios. The ratios will be like-for-like replacements of Altman's ratios, i.e. a liquidity ratio for a liquidity ratio etc.

The next section presents our data and sample selection. The models are based on Norwegian data of both listed and unlisted firms from 2009 to 2016. All industries except finance and insurance are included in the study.

This is followed by a presentation and discussion of the results and an evaluation of the model. Finally, we present the conclusions drawn from this study and suggest future research on the topic.

2. Literature review

2.1 Previous research on bankruptcy prediction

The main focus of this study is Altman's Z´´-model. However, there has been extensive research of the topic of bankruptcy prediction over time. This sub-section serves as an overview of past research on the topic.

The pioneering work on bankruptcy prediction was done by William Beaver (1966). His univariate analysis set the stage for the development of future models. His sample consisted of 79 firms from different industries, which had failed in the years 1954-1964 (Beaver, 1966). The bankrupt firms were matched with non-bankrupt firms on the basis of asset size and industry (Beaver, 1966). Beaver computed 30 ratios and concluded that the cash flow/total debt ratio was the best ratio. The ratios could discriminate between bankrupt and non-bankrupt firms up to five years prior to bankruptcy. However, the error rate increased as the time-span prior to bankruptcy increased (Beaver, 1966).

In 1968, Altman developed his Z-score model using MDA. This is one of the most well-known and applied models. A thorough presentation of this model will be given separately as the related Z´´-model is the foundation of this paper.

Altman, Haldeman and Narayan (1977) built upon earlier Z-score models and developed the ZETA model (Altman E. , 2000). The sample consisted of 54 bankrupt and 58 non-bankrupt firms from 1967-1975. The ZETA™ Model included seven ratios, an increase compared to Altman´s earlier models. However, because the model is a proprietary effort, the coefficients are undisclosed. The new model was able to classify bankrupt companies up to five years prior to failure, with a 90% success rate one year prior to bankruptcy and 70% success rate five years prior.

Other research related to the Z-score model includes Taffler (1983), who adapted the model to UK firms. Deakin (1972) combined Beaver´s ratios with Altman´s methodology, trying to find the best linear combination of the ratios from Beaver´s study (Altman E. , 1983).

Ohlson (1980) used logistic regression to predict bankruptcy. He argued that MDA had several problems that made it an inferior approach. First, the statistical requirements imposed on the variables were difficult to satisfy. Furthermore, he argued that the discriminant score had no intuitive interpretation, and that the matching procedure of bankrupt and non-bankrupt firms was questionable.

Ohlson gathered financial data from 105 bankrupt and 2058 non-bankrupt firms from the years 1970-1976. This was a considerable increase from earlier research, and the ratio of bankrupt to non-bankrupt firms was more representative to the actual ratio. An important finding was that the analysis was sensitive to when financial data is made available to the public. Because firms in distress are more exposed to inaccurate accounting, Ohlson only included information that was available prior to the bankruptcy.

Throughout the years, much emphasis has been put on methodology. Zmijewski (1984) used a probit model to predict bankruptcy. He developed a model with three financial ratios. The sample consisted of 40 bankrupt and 800 non-bankrupt industrial firms, from the years 1972-1978.

Zmijewski was particularly concerned with the sampling process of earlier research, and pointed out two problems. The first problem arises when researchers match the samples of bankrupt and non-bankrupt firms. When the non-bankrupt firms are chosen based on the characteristics of the bankrupt firms, this is no longer random sampling. Second, a common problem with data sets are missing values, forcing the researcher to drop observations. As a result, the researcher needs to assume that the observations dropped were a representable ratio of bankrupt and non-bankrupt firms. However, this may not be the case as the quality of the financial statement is likely to be lower for distressed firms.

Another approach was suggested by Shumway (2001). He used hazard models in his analysis. This approach had the advantage of using all available information, spanning over several years (Shumway, 2001). By contrast, the static logit model can only use one year for each observation. He argued that financial ratios change considerably from year to year, making static models inappropriate for bankruptcy prediction.

In 2004, Hillegeist et al. developed a model utilizing the insight of the Black-Scholes option-pricing model (Hillegeist, Keating, Cram, & Lundstedt, 2004). This model was heavily based on market values. The performance of the model was tested, and Hillegeist et al. concluded that their model significantly outperformed Altman and Ohlson´s accounting-based models.

In Norway, Norges Bank use the SEBRA model to estimate bankruptcy probabilities for Norwegian limited companies (Bernhardsen & Larsen, 2007). It is also used to estimate the expected losses on loans to firms. The model was developed using a large database of Norwegian firms from the years 1990-1999. In 2007, the model was revised and two new models, the SEBRA Basic and SEBRA Extended, were introduced. The two models were a simplification and a refinement of the original model. After testing the new models, they found that the simple SEBRA model only had a marginally lower accuracy rate for bankruptcy prediction.

In recent times, following technological development, neural network (NN) analysis has emerged as an alternative approach (Caouette, 2008). A neural network is a collection of simple, interconnected computational elements. The computer identifies and learns links and patterns between the data units, and use it to solve given problems. There are several studies applying NN. Charitou applied a NN analysis on a sample of UK firms from the years 1988-1997, concluding that NN analysis provide at least as good results as the more traditional methods (Charitou, Neophytou, & Charalambous, 2004).

Gissel et al. (2007) reviewed a large number of bankruptcy prediction models from 1930 to 2007. The paper summarizes existing research on bankruptcy prediction studies, taking into account 165 different studies. It focuses on how bankruptcy prediction studies have evolved, both in terms of different methods, the variety of ratios and its applications. The study also provides an overview of the most used financial ratios.

In general, the study found that discriminant analysis was a popular approach between the 1960's and 1980's. However, logit analysis became more popular throughout the 1980's and 1990's. This was followed by the emergence of neural networks in the 1990's which is still popular today.

2.2 Review of Altman's Z-score models

The Z-score model has been the subject of several studies over time. Although it is widely regarded as a successful model, it has been evaluated and criticized by researchers. Based on the revisions Altman made to his original model, it is clear that he was aware of at least some of the limitations with the Z-score. It is important to note that although we criticize the model, it is not necessarily the case that Altman had the opportunity to produce a better model, given the limited information disclosed in annual reports at the time.

In order to evaluate the effectiveness of a model, it is important to test it using hold-out samples. Altman did test his original model with a hold-out sample. When testing with only bankrupt firms, the model proved to be very accurate, classifying 96% of the bankrupt firms correctly. The secondary sample of only non-bankrupt firms contained firms under financial distress, but despite this it correctly classified 79% of the firms. This gives an overall accuracy of 83.5% for the out-of-sample firms, albeit using distressed non-bankrupt firms. Altman did not perform any hold-out sample tests on his revised Z´- and Z´´-model. This was due to the lack of data for private firms.

We have compiled a collection of 16 different studies that have used one of Altman's Z-score models to evaluate the default risk of firms (Appendix 1). The collection consists of studies from different time periods and different countries, using both Altman's original coefficients and re-estimated coefficients. These studies use hold-out-samples that test the generalizability of the model. This collection is the basis for our review of Altman's models.

Although our study focuses on the Z´´-model, the review includes research on all three models. This is because there is limited research on the Z´´-model. However, the models are relatively similar, and the criticisms are generally relevant for all of them.

Note that the discussion below is based on a sample of 16 different studies. The results may therefore be affected by random factors and sample bias. Nevertheless, we think that the general arguments made below still hold, even if the arguments should be interpreted with caution.

Charles Moyer (1977) performed one of the first reviews of the Z-score model. He emphasized that the hold-out sample tests conducted by Altman were done using the same time period as the estimation sample. Because the model should predict bankruptcies in the future, the argued that the model should be tested with a hold-out sample from a later time period.

Moyer used a sample of firms from the years 1965-1975. It included 27 bankrupt and 27 non-bankrupt firms. However, some firms were later dropped. The asset size ranged from \$15 million to \$1 billion, which was higher than for the sample employed by Altman.

The original model had an overall accuracy of 75%, with the Type 1 and Type 2 error rates at 39.2% and 12%, using Moyer's sample. The accuracy is significantly different from Altman's hold-out-sample. This is particularly true for the prediction of bankrupt firms, which was 35.2% lower.

Moyer developed two re-estimated models. The first model was a simple re-estimation of the coefficients, which yielded an improved accuracy rate at 88.1%. However, the test was done using the same sample that was the basis for the re-estimation. This is conflicting with Moyer´s own argument that the model should be tested on a secondary sample.

A second, re-estimated model was developed using stepwise estimation. This procedure test all possible combinations of the variables, and yields the model where the Wilks´ lambda is minimized. This re-estimation only included three out of five variables.

The accuracy with this model was 90.5%. The Type 1 error rate was the same, while the Type 2 error rate was reduced to 14%.

Table 1: Results from Moyer (1977)

Sample	Model	Overall	Bankrupt	Non-bankrupt
Altman hold-out	Original	83.5% (91)	96% (25)	78.8% (66)
1965-1975	Original	75% (48)	61% (23)	88% (25)
1965-1975	Re-estimated	88.1% (42)	95% (20)	82% (22)
1965-1975	Re-estimated, stepwise	90.5% (42)	95% (20)	86% (22)

Numbers in parenthesis represent the total number of firms within the group

Grice and Ingram (2001) wanted to check the generalizability of the Z-score model by testing if it could produce accurate predictions for more recent data. Furthermore, they wanted to test if the Z-score model could predict bankruptcies for non-manufacturing firms as accurately as for manufacturers.

The study included two different samples, an estimation sample and a hold-out sample. The estimation sample was used to re-estimate the coefficients of the Z-score. This included 141 distressed and 824 non-distressed firms from 1985-1987. The hold-out sample consisted of 148 distressed and 854 non-distressed firms from 1988-1991. As opposed to Altman´s sample, both samples included firms from a wide range of industries. The approach, using an estimation sample and a newer hold-out sample, is in accordance with Moyer’s reasoning.

The accuracy of the original Z-score model, when applied on the hold-out sample, was significantly lower than Altman´s tests. The overall accuracy was 56.1%, compared to 83.5% for Altman´s hold-out sample tests (Grice & Ingram, 2001). This indicates that the original Z-score model is not as accurate for predicting bankruptcies in recent times. However, the hold-out sample contained industries not intended for the original model.

A subsample of the hold-out sample, only including manufacturing firms, was also tested. The overall accuracy for the manufacturing sample was 69.1%. This was still lower than the 83.5% rate for Altman. The accuracy for manufacturing firms was significantly higher than for the overall sample, suggesting that the model is better for its intended industry.

Next, the researchers used the estimation sample to re-estimate the coefficients. There were significant differences between the original and re-estimated coefficients. The coefficients are presented in Table 2. Limiting the estimation sample to only manufacturers also had an impact on the re-estimated coefficients. The re-estimated coefficients are negative for several ratios, but this is not commented by the researchers.

Table 2: Coefficients from Grice & Ingram (2001)

Model	**X_1**	**X_2**	**X_3**	**X_4**	**X_5**
Z-score	1.200	1.400	3.300	0.600	0.990
Re-estimated	0.831	1.504	2.073	-0.014	-0.058
Re-estimated, manufacturing	-0.386	2.067	1.385	-0.005	-0.069

The overall accuracy using the re-estimated coefficients was 87.6%, significantly higher than the 57.0% produced with the original coefficients. When applied on the sample with only manufacturing firms, the overall accuracy was 86.4%. Compared to Altman´s test, the overall accuracy when re-estimating was higher. However, the accuracy for bankrupt firms was 47.4% lower compared to Altman´s hold-out test.

Table 3: Results from Grice & Ingram (2001)

Sample	**Model**	**Overall**	**Bankrupt**	**Non-Bankrupt**
Altman hold-out	Original	83.5% (91)	96.0% (25)	78.8% (66)
1988-1991	Original	56.1% (972)	68.2% (85)	54.9% (887)
1988-1991: Manufacturing	Original	69.1% (547)	69.2% (78)	69.1% (469)
1988-1991	Re-estimated	87.6% (972)	48.6% (85)	94.9% (887)
1988-1991: Manufacturing	Re-estimated	86.4% (547)	55.4% (78)	92.1% (469)

Numbers in parenthesis represent the total number of firms within the group

Begley et al. (1996) performed a similar study. They tested the original model, as well as re-estimating the coefficients. The hold-out sample included 65 bankrupt and 1300 non-bankrupt firms from 1980-1989. All the firms were listed on stock exchanges in the US, and represented a wide range of industries.

Compared to Altman´s test, the performance is less accurate. The overall accuracy fell from 83.5% in Altman´s test to 78.2% with the more recent sample. The Type 1 error rate is significantly higher at 18.5%, while the Type 2 error rate is similar at 25.1%.

As opposed to Grice and Ingram, Begley et al. found that re-estimating the model did not significantly change the results. This is inconsistent with past results, particularly because the accuracy is based on the estimation sample. The re-estimated model was not tested on a separate hold-out sample. The overall accuracy increased by 0.2 percentage points, while the Type 1 error rate increased to 21.5%. Given that a Type 1 error is costlier, the researchers argue that the original model was preferred.

Table 4: Results from Begley et al. (1996)

Sample	Model	Overall	Bankrupt	Non-bankrupt
1980-1989	Original	78,2 (1365)	81,5 (65)	74,6 (1300)
1980-1989	Re-estimated	78,4 (1365)	78,5 (65)	78,4 (1300)

Numbers in parenthesis represent the total number of firms within the group

In more recent times, Gutzeit and Yozzo (2011) have reviewed the model. Their study focused on the original model´s performance during the most recent recession in 2007-2008. The duration and severity of the recession, causing a large number of bankruptcies in the US, generated a sufficiently large set of data to be tested.

The study limited the sample to large, publicly owned manufacturing firms with total assets or sales in excess of $50 million in 2007. The researchers acknowledged that the economy had become highly service-intensive in the last decade, making the model less relevant. However, they limited the sample to manufacturing firms to comply with Altman´s limitations.

Using the sample from the recession, the accuracy for correctly classifying bankrupt firms was 90% one year prior to bankruptcy. The high accuracy rate is attributed to the inclusion of the market value of equity. The researchers found that it accounted for 40-50% of a typical non-bankrupt firm´s Z-score, but only 10-20% for a bankrupt firm´s Z-score.

The importance of the market variable was further underlined when the same sample was applied to the revised Z´-model, which only include book values. The accuracy fell from 90% to 75% for the bankrupt firms one year prior to bankruptcy, and from 69% to 58% two years prior.

The models were also tested for a sample of non-bankrupt firms. The Z-scores were computed for every year from 2004 to 2008. The Type 2 error rate increased during the recession, peaking at 29.8% in 2008. These results were consistent with the significant drop in market values of equity during the recession (Gutzeit & Yozzo, 2011).

There has been a considerable effort to review and adjust the model in an international environment. The model has been tested in in several countries throughout the world. We have gathered results from 11 studies, which are summarized in Table 5. These studies are based on data from different countries and time periods, and show some general trends. The numbers presented below include studies using Altman's original coefficients as well as re-estimated coefficients.

Table 5: Summary of results from studies outside the U.S

Study	Country	Years	Overall	Bankrupt	Non-bankrupt
Almamy, Aston & Ngwa (2016)*	UK	2000-2013	54.4%	60.6%	54.0%
Jackson & Wood (2013)*	UK	2000-2009	40.1%	52.0%	39.9%
Jeroen Avenhuis (2013)d*	Netherland	2008-2012	80.6%	35.7%	82.5%
Bruno, Keglevic, Tanja (2014)c	Croatia	2008-2011	80.0%	70.0%	90.0%
O. Machek (2014)	Czech Rep.	2007-2012	44.3%	-	-
Celli (2015)	Italy	1995-2013	87.3%	84.3%	90.1%
Christopoulos, Gerantonis & Vergos (2009)	Greece	2003-2007	56.6%	65.9%	54.2%
Wang & Campbell (2010)	China	1998-2008	51.2%	96.3%	51.1%
Wang & Campbell (2010)*	China	1998-2008	84.7%	85.2%	84.7%
Bandyopadhyay (2006)*	India	1998-2003	83%	82%	84%
Pongsatat, Ramage & Lawrence (2004)a	Thailand	1998-2003	58.9%	90.5%	40.0%
Pongsatat, Ramage & Lawrence (2004)b	Thailand	1998-2003	64.1%	94.9%	16.0%
Lifschutz (2010)	Israel	2000-2007	62.5%	100%	25%

(*) indicates that the coefficients are re-estimated, (a) indicates only large asset firms, (b) indicates only small asset firms, (c) indicates small sample size, (d) indicates master thesis

First, we note that the accuracy of the studies differs greatly. The overall accuracy ranges from 40.1% to 87.3%. In addition, we see that the prediction accuracy for bankrupt and non-bankrupt firms fluctuates to a large extent. These results are very different from the results in Altman's study.

In general, we see that the overall accuracy is slightly weaker for studies performed outside the U.S. Furthermore, the variation in overall accuracy is also larger for studies performed outside the U.S. We also see that the results vary greatly within relatively similar geographical areas. In Europe, the accuracy ranges from 40.1% to 87.3%. In Asia, the accuracy ranges from 51.2% to 84.7%. Again, we find that re-estimating the coefficients improves accuracy.

There are differences between studies conducted in different geographical areas. If we compare the overall accuracy for the European and Asian (excluding Israel) samples, we see that the accuracy is relatively similar, with 63.3% for the former and 68.4% for the latter. However, if we compare the accuracy for bankrupt firms, the accuracy is significantly higher for Asian countries at 89.8% compared to 64.0% for European countries.

Regardless of country, re-estimating the coefficients seems to improve the results. This is true both when we compare results within each study, and when comparing the average overall accuracy of the re-estimated and original coefficients. We find that the increase in overall accuracy, when re-estimating, is caused by a decrease in Type 2 errors. This is consistent with recent statements by Altman, who acknowledged that the original model has been producing more Type 2 errors in more recent samples (Altman E. , 2000). This issue is partially resolved by re-estimating the coefficients. However, re-estimating the coefficients generally increases the Type 1 error. For three out of four studies, re-estimation results in a reduced accuracy for bankrupt firms.

In general, we see that the overall accuracy is lower than that of the hold-out-sample in Altman's study. This is the case when using both the original and re-estimated coefficients. Additionally, we see that the accuracy with regards to bankrupt firms is far from the 96% achieved in the original study.

There are several potential explanations for these findings. We generalize these into two categories: temporal differences and geographical differences. Temporal differences refer to differences between the business environment when Altman performed his original study and later studies. These are relevant for both the U.S. and international studies. Geographical differences refer to differences in the business environment caused by the geographic affiliation of the firms. These are most relevant for the international studies, but there are also potential differences inside a large economy like that of the U.S.

The most important evidence for temporal differences is the effect re-estimating the coefficients has on the accuracy of the model (Grice & Ingram, 2001). We also see that the accuracy of newer studies is lower than that of older studies. If we consider studies based on U.S. data we generally see that re-estimation produces significant changes to the overall accuracy. They also show that the overall accuracy decreases as the samples are derived from years further away from Altman's original sample. This is an important insight as the studies performed using a sample of U.S. firms keep the geographical aspect relatively constant.

Following these findings, we need to discuss the reason for why there are temporal differences. These can be general changes in the business environment, such as an overall increase in default rates. Gentry, Newbold and Whitford (1985) found that that the average business failure rate ranged from 0.38% to 1.19% between 1970 and 1991. One reason for this might be that competition has increased. Hence, there might be different requirements today for profitability, solvency, liquidity and the like. This might affect the general accuracy of older bankruptcy prediction models as the ratios included measure these aspects. The coefficients and cut-off points are thus designed with different requirements in mind.

Temporal differences can also be explained by specific changes that affect the ratios included in the model. Sherbo and Smith (2013) emphasize the growth of the economy and that the market capitalization of the S&P 500 was 32 times higher in 2013 than in 1968. This is especially important for the market-based variable in the original model. Sherbo and Smith claim that market values have increased substantially more that accounting values.

The aforementioned market aspect is not present in the Z´´-model as it only includes book values. However, we argue that the relative increase in market values compared to book values might indirectly affect book values. One example of this is goodwill, which is the residual of the purchase price and fair market value. If the premium is constant over time, e.g. 5%, the absolute value of goodwill will increase following an increase in market values. This argument would be consistent with research that shows that the goodwill to assets ratio has increased over time (O'Shaughnessy, 2015).

Altman used total assets to adjust for firm size, on four out of five ratios. This was an appropriate measure of size in the 1960s as companies were more homogenous and most companies were asset heavy manufacturers. Over time, companies have become more heterogeneous. The relationship between asset size and sales is different than in the 1960s. This is evident from the fact that the change in total assets over time is different from the change in sales over time (Yardeni, Abbott, & Quintana, 2017).

There has also been a change in the composition of the economy. This is evident from the emergence of companies with lean balance sheets but large sales incomes, such as service companies (US Bureau of Labour Statistics). These companies are different from asset heavy manufacturing firms. This represent a problem with using total assets to adjust for size, when sales would be more appropriate for asset light companies.

Changes in accounting practices over time also have an effect on the predictive ability of the model. One major change is the introduction of International Financial Reporting Standards (IFRS). More companies are moving over to using IFRS when reporting financial data (Deloitte).

This represents a structural difference between when Altman performed his analysis and today. The use of historical cost was, for instance, more prevalent in the 1960s than today. Today firms use alternative valuations methods for all types of items. This and other differences can affect the predictive ability of the model.

We also need to account for geographical differences. Naturally, the business environment differs between countries and it is important to keep in mind that Altman's model was based on U.S. data. The aforementioned studies show differences between countries, where the model tends to be more accurate in the U.S.

We argue that there are several reasons for these differences. Different accounting practices between countries can arguably explain some of the differences. These differences can affect how items in the financial statements are estimated, which naturally reduces the accuracy of the model when comparing firms from different countries.

Furthermore, the company default rate varies between countries. Companies in some countries may therefore be exposed to a greater default risk. This can for instance be driven by government policies, economic stability, and competition. One example is how severely the financial crisis affected the EU compared to Norway (Eurostat, 2017). This arguably resulted in an increased default risk in the EU compared to Norway.

Altman´s studies have been subject to extensive research. In his article on international bankruptcy prediction models, Altman provides an overview of models from several different countries (Altman E. , 1984). It covers models from Japan, West Germany, Brazil, Australia, England, Canada, Netherlands, and France.

These models have a similar approach as Altman, i.e. MDA with a set of financial ratios as predictors. Therefore they might be seen as refinements of the Z-score model. However, we consider these models to be independent models. This is because these studies differ from Altman´s model with regards to the financial ratios included, both in the number of ratios and the characteristics they are describing.

Some of the studies are more similar to the Z-score models. This is particularly true for the Canadian and Brazilian models, where Altman was involved in the studies. The following table summarize the models presented in the article.

Table 6: Studies using models similar to the Z-models					
Study	**Country**	**Years**	**Number of ratios**	**Type of firms**	**Overall accuracy**
Takahashi et al. (1979)	Japan	1962-1976	8	Manufacturing, listed	81.2%
Ko (1982)	Japan	1960-1980	5	Manufacturing, listed	82.9%
Weinrich (1978)	West-Germany	1969-1975	6	-	89% (two years prior)
Altman, Baidya & Ribeiro-Dias (1979)	Brazil	1975-1977	5	Mixed, listed	88%
Castagna & Matolcsy (1981)	Australia	1963-1977	10	Manufacturing, listed	-
Taffler & Tisshaw (1977)	England	1969-1975	4	Manufacturing, listed	97%
Altman & Lavallee (1981)	Canada	1970-1979	5	Manufacturing & retail, listed	83.3%
Van Fredrikslust (1978)	Netherlands	1954-1974	2	Mixed, listed	92.5%

In general, the overall accuracy is higher for these models than for the Z-score models when they are used on hold-out-samples. This is expected as these models are specifically designed for the sample and time they evaluate.

The focus of our study lays closer to the Z´´-model. We have not been able to find any studies that have the same approach as this study. Nor have we been able to find reviews of the ratios which is the focus of this study. One of the reasons why we have found limited studies evaluating and improving the Z-score models is that the focus has been on developing new models using new approaches. Therefore, none of the studies mentioned in this section are directly applicable to our study.

Nevertheless, the results from these studies provide some insight that is important to consider. We see that prediction accuracy for older models decreases over time. Furthermore, we find that newer studies, using re-estimated coefficients, struggle to produce results as accurate as Altman's original study. This is arguably because the economy today is more complex and heterogeneous. Hence, it is more difficult to develop a general model for bankruptcy prediction. Based on our interpretation of these findings, we argue that specific models are superior and should be applied if possible.

3. Description of MDA

In this section, the multivariate discriminant analysis that forms the basis for the study is discussed. Additionally, four methods for evaluating the models are presented, namely classification matrices, Wilks' Lambda, receiver operating curves, and McNemar´s test. Lastly, we describe four statistical tests that assess whether the models satisfy the underlying assumptions.

3.1 Multivariate Discriminant Analysis

MDA is a statistical tool used to study the differences between two or more groups of objects, with respect to multiple variables simultaneously (Klecka, 1980). It was introduced by Fisher (1936) when he proposed a technique that maximized the group differences, while minimizing the variation within the groups.

In order to apply MDA, there are several prerequisites and assumptions (Klecka, 1980). It requires two or more mutually exclusive groups. The groups must be defined so that each observation only belongs to one group. The number of independent variables cannot exceed n – 2. There must also be at least two observations in each group. Furthermore, the researcher must be able to discriminate between the groups on the basis of a set of characteristics. These are called discriminating variables, and must be measured on an interval or ratio level.

There are several limitations to the statistical properties of the discriminating variables. First, a variable cannot be a linear combination of any other variable. Likewise, two variables that are perfectly correlated cannot be included.

Second, the population covariance matrices need to be relatively equal for each group. The MDA in this study employs a linear discriminant function, which is a simple linear combination of the discriminating variables. The assumption of equal group covariance matrices allows for simplification of the procedure of deriving the coefficients, as well as allowing for tests of significance (Klecka, 1980).

Third, MDA assumes that each variable is normally distributed when drawn from the population. This assumption permits precise computations of tests of significance and probabilities of group significance. It also assumes multivariate normality, meaning that the group is drawn from a population with a multivariate normal distribution on of the discriminating variables (Klecka, 1980).

The assumptions of MDA are relatively strict, and it is stated as one of the most important reason for choosing other approaches (Ohlson, 1980). If the data do not satisfy the assumptions, the statistical results will not be a precise reflection of reality. Nevertheless, several researchers have found that MDA is a rather robust technique that can tolerate some deviations from the assumptions (Klecka, 1980). However, the method is very sensitive to outliers.

It is difficult to determine how much deviation the model can tolerate. However, if one is mainly interested in a model that can predict well or describe the real world, the accuracy is the most important factor to consider (Klecka, 1980).

The space dimensionality in discriminant analysis equals the number of groups minus one (Altman E. , 2000). Because the number of groups in this study equals two, the analysis is transformed into a simple, one-dimension analysis. The discriminant function looks like a regression analysis on the surface. However, the mathematical functions used to derive the coefficients are different. In its simplest form, the MDA has the following form:

$$Z = V_0 + V_1X_1 + V_2X_2 + \cdots + V_nX_n$$

$$V_1, V_2, \ldots, V_n = raw\ coefficients$$
$$X_1, X_2, \ldots, X_n = independent\ variables$$

The discriminant coefficients are derived so that a linear combination of the variables maximize the difference between the groups (Altman E. , 1968). The coefficients yielded by the function are considered raw coefficients. They are useful for classification purposes, but the scores they produce have no obvious meaning.

However, a simple adjustment to the values give the coefficients an explanatory value. These latter coefficients are defined as:

$$U_i = V_i\sqrt{N-g} \text{ and } U_0 = -\sum_{i=1}^{p} U_iX_i$$

$U_i = discriminant\ coefficients\ on\ standard\ form$

$V_i = raw\ coefficients$

$N = total\ number\ of\ cases\ over\ all\ groups$

$g = number\ of\ groups$

The coefficients on standard form causes the discriminant scores, over all cases, to have a mean of zero and within-groups standard deviation of one (Klecka, 1980). The transformation means that each axis is stretched or shrunk such that the score represents the number of standard deviations it is from the overall mean. This means that the user can immediately understand the relative score, and if this score is high or low.

3.2 Methods for evaluating results from MDA

3.2.1 Classification matrix

The prediction accuracy of the models will be presented using a classification matrix. The accuracy is a plain and simple method of assessing the model, but it is also an important one because the ultimate goal is to predict correctly. There are two different types of error presented in the matrix. The classification of a bankrupt firm as non-bankrupt (Type 1 error) and the classification of a non-bankrupt firm as bankrupt (Type 2 error).

There are different costs associated with the different types of error. In the context of bankruptcy prediction, costs associated with Type 1 errors are a bank´s loss of principal and interest or an investor´s loss of investment. For Type 2 errors, costs are forgone profit because of avoiding investment opportunities.

Table 7: Outline of a Classification Matrix

		Observed	
		Bankrupt	Non-bankrupt
Classified	Bankrupt	Correct	Type 2 error
	Non-bankrupt	Type 1 error	Correct

3.2.2 Wilks' Lambda

Wilks´s lambda is a measure of the overall significance of the model (Klecka, 1980). The test proceeds indirectly, meaning that rather than testing the function itself we test the residual discrimination in the system prior to deriving the actual function.

Wilks´s lambda is defined as the ratio of within-groups sums of squares to the total sums of squares (Stevens, 2009). This is the proportion of the total variance in the discriminant scores not explained by differences among groups. If the statistic takes a value near zero it means high discrimination. In this situation, the group means are greatly separated and very distinct relative to the amount of variance within the group (Klecka, 1980). On the other side, if the value equals 1.0 there are no differences between the group means.

3.2.3 ROC Curve

A Receiver Operating Characteristics (ROC) graph is a technique for visualizing, organizing, and selecting classifiers based on their performance (Fawcett, 2006). It is applicable to binary classifier systems, e.g. a model that predicts bankruptcy/non-bankruptcy. Every outcome will be classified in accordance with the matrix in Table 8.

Table 8: Outline of a Contingency Table for ROC curve

		True condition	
		Condition positive	Condition negative
Predicted condition	Predicted condition positive	True positive	False positive
	Predicted condition negative	False negative	True negative

In our study, a bankruptcy is defined as a "positive" while a non-bankruptcy is defined as a "negative". A true positive is a correct classification of a positive, while a true negative is a correct classification of a negative. Correspondingly, a false positive is a predicted positive with a true negative outcome. A false negative is a predicted negative with a true positive outcome.

ROC graphs are two-dimensional, where the true positive rate is plotted on the Y-axis and the false positive rate is plotted on the X-axis. The rates are calculated as follows:

$$True\ positive\ rate = \frac{True\ positives}{Sum\ of\ true\ positives\ and\ false\ positives}$$

$$False\ positive\ rate = \frac{False\ positives}{Sum\ of\ true\ negatives\ and\ false\ negatives}$$

In order to compare different classifiers, the area under the ROC (AUROC) graph is a common measure. The AUROC is equivalent to the probability that the classifier will rank a randomly chosen positive higher than a randomly chosen negative, assuming that a positive ranks higher than a negative (Fawcett, 2006). The area can range from 0.5 to 1.0. A random classifier will on average have a value of 0.5 while a perfect classifier will have a value of 1.0. An AUROC value of 0.7-0.8 shows acceptable discrimination, 0.8-0.9 shows excellent discriminations and higher than 0.9 shows outstanding discrimination (Hosmer & Lemeshow, 2000).

3.2.4 McNemar's Test

In order to assess and compare the performance of the different models, we use a test introduced by Quinn McNemar (1947). The test exists in several versions (Fagerland, Lydersen, & Laake, 2013). In this paper, we will use the exact binominal version, due to the size of the sample. For bigger samples the asymptotic version, based on the normal distribution, can be used.

The joint performance of the classification methods can be summarized in a contingency table as follows:

Table 9: Outline of a Contingency Table Showing Classifications of Two Models

		Method 2		
		Correct	Incorrect	Sum
Method 1	Correct	n_{11}	n_{12}	n_{1+}
	Incorrect	n_{21}	n_{22}	n_{2+}

In this framework, n_{11} represents the number of correct classifications for both method 1 and 2, while n_{22} is the number of incorrect classifications for both methods. n_{12} is the number of classifications where only method 1 is correct, while n_{21} represents the opposite. The notation for the outcome probabilities p_{ij} follow the same layout.

We define the probability that that method 1 is correct as p_{1+} and the probability that method 2 is correct as p_{2+}. The null and alternative hypothesis of interest are defined as:

$$H_0: p_{1+} = p_{2+}$$
$$H_1: p_{1+} \neq p_{2+}$$

Because the number of misclassifications in the different models are related, we cannot test the misclassification directly (Næss, 2015). McNemar´s test offers a solution to this problem by only considering the discordant pairs. The discordant pairs are defined as the cases where the classifications differ, i.e. $n_{12} + n_{21}$ from the matrix.

The test statistic measures the strength of evidence against the null hypothesis. We use n_{12}, conditional on the number of discordant pairs $n = n_{12} + n_{21}$ as a simple test statistic (Fagerland, Lydersen, & Laake, 2013). The conditional probability under the null hypothesis of observing any outcome x_{12}, given n discordant pairs, is the point probability:

$$f(x_{12}|n) = \binom{n}{x_{12}}\left(\frac{1}{2}\right)^n$$

The McNemar´s exact conditional two-sided p-value is obtained by:

$$Two\ sided\ p - value = 2 \sum_{x_{12}=0}^{\min(n_{12},n_{21})} f(x_{12}|n)$$

The p-value equals the probability of observing the observed values, or more extreme values, when the null hypothesis is true. We reject the null hypothesis if the p-value is lower than α. In this case we conclude that there is a significant difference between the methods.

3.3 Methods used to test the assumptions of MDA

3.3.1 F-test

In this study, an F-test is used to determine whether the group means for bankrupt and non-bankrupt firms are equal for a given ratio. An F-test is a statistical test where the distribution of the test statistic follows an F-distribution. If the F-statistic is greater than the criteria, we conclude that the means are significantly different. The F-statistic is given by the following ratio:

$$F = \frac{variation\ between\ sample\ means}{variation\ within\ samples}$$

$$F = \frac{MS_{Groups}}{MS_{Error}} = \frac{\frac{SS_{Groups}}{(G-1)}}{\frac{SS_{Error}}{n-G}}$$

In the equation above, MS is mean square, SS is sum of squares, G is number of groups and n is number of cases. The F-statistic is compared with the F-distribution with numerator degrees of freedom equal to G – 1 and denominator degrees of freedom n – G.

3.3.2 Chi-square difference test

We use a chi-square difference test to evaluate whether a model is statistically different from a nested model, i.e. a model containing one less variable. The test takes into account the difference in chi-values and differences in degrees of freedom (df).

$$\chi^2_{difference} = \chi^2_{nested} - \chi^2$$
$$df_{difference} = df_{nested} - df$$

In order to test whether the difference is significant we use a chi-table. We use the chi-difference value and the difference in df in the chi-table. If the difference is significant, we can conclude that the unrestricted model is different from the nested model (Schermelleh-Engel & Werner, 2010).

3.3.3 Shapiro-Wilk test

The Shapiro-Wilk test is used to test for normality, on an individual basis, in the different ratios applied. The null hypothesis is that the ratio is normally distributed. Hence, a rejection of the null hypothesis indicates that ratio is non-normal. The Shapiro-Wilk test is considered to be conservative and often wrongly rejects normality, especially in large samples (Field, 2009). We also check for normality using histograms of frequency and normal Q-Q plots. The Shapiro-Wilk test statistic can be estimated using the following formula:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$(a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{\frac{1}{2}}}$$
$$m = (m_1, \dots, m_n)^T$$

Where $x_{(i)}$ is the nth-smallest observation in the sample and $(m_1, \dots, m_n)$ is the expected value of the kth-observation, ordered from smallest to largest, from the independent and identically distributed (i.i.d.) variables. The statistic W is tested against a critical value from a Shapiro-Wilk table, using α significance level and n observations.

3.3.4 Box's M test

The Box's M test is a statistical test that considers whether covariance matrices are homogenous. The null hypothesis is that there is homogeneity, hence rejecting it indicates a violation of one of the assumptions for MDA. The test is however conservative and log-determinants are better to use for large samples (Manly, 2004). A thorough explanation of the test is outside the scope of this study and is therefore not included.

4. Description of Altman's Z-score models

Altman found earlier research using univariate analysis to be inconclusive (Altman E. , 1968). As a result, he developed a model using MDA. The technique had the advantage of considering the entire set of characteristics of a firm and the interactions of these.

Altman started by sampling a group of bankrupt firms and a corresponding group of non-bankrupt firms (Altman E. , 1968). The original sample consisted of 66 firm, with 33 firms in each group. The bankrupt firms where all manufacturing firms that filed a bankruptcy petition during the period 1945 – 1965. The non-bankrupt firms were paired on a stratified random basis, with an asset size range restricted between $1-25 million. The mean asset size of non-bankrupt firms was $9.6 million, compared to $6.4 million for the bankrupt firms. The non-bankrupt firms were all in existence 1966.

After the sample was selected, the income statements and balance sheets were collected. Based on these, Altman calculated 22 potential financial ratios he considered to be helpful predictors. The potential ratios were chosen on the basis of their popularity in the literature and their relevance to bankruptcy prediction. The ratios were categorized into five classes (Altman E. , 2000):

Liquidity:

1. Current ratio
2. Cash and marketable securities/Current liabilities
3. Current assets – Current liabilities/Total assets

Profitability:

4. Gross profit/Sales
5. Profit before taxes/Sales
6. Profit after taxes/Sales
7. Profit after taxes before interest/Total assets
8. Profit before taxes and interest/Total assets
9. Number of years of negative profits in last 3 years

Leverage:

10. Short term debt/Total assets
11. Long term debt/Total assets
12. Total debt/Total assets

Solvency:

13. Retained earnings/Total assets
14. Market value of equity/Par value of debt
15. Net worth/Total debt

Activity:

16. Sales/Cash and marketable securities
17. Sales/Inventory
18. Cost of goods sold/Inventory
19. Sales/Net fixed assets
20. Sales/Current liabilities
21. Sales/Total assets
22. Working capital/Sales

Five ratios were chosen as doing the best job of predicting bankruptcy. Because many financial ratios are highly correlated with each other, the number of ratios utilized are limited. The final profile was selected on the basis of an overall decision process. Multiple ratio profiles were tested, observing the statistical significance of the model and the relative contribution of the independent variables. Furthermore, the intercorrelations between the variables, the predictive accuracy and Altman´s own judgment was considered.

The final model did not include the most significant ratio measured individually. However, because of correlations between the ratios, the overall accuracy was better with the selected ratios. Compared to univariate analysis this was a major development. The discriminant function of the Z-score model is as follows:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5$$

$X_1 =$ Working Capital / Total Assets
$X_2 =$ Retained Earnings / Total Assets
$X_3 =$ Earnings Before Interest and Taxes / Total Assets
$X_4 =$ Market Value of Equity / Book Value of Total Debt
$X_5 =$ Sales / Total Assets

Working capital/Total assets: This is a measure of the net liquid assets relative to the firm´s total capitalization. Working capital is defined as current assets minus current liabilities. A firm with operating losses will normally experience shrinking current assets in relation to total assets, resulting in a lower ratio.

Retained earnings/Total assets: Retained earnings is a measure of profitability over time. This measure favors older firms because profitability over time is likely to translate into a high ratio of retained earnings. Since empirical studies show that young firms are more likely to go bankrupt, the ratio does not unfairly discriminate against young firms.

EBIT/Total assets: This is a measure of the productivity of the firm, leaving out any tax or leverage factors. Ultimately, a firm's existence is based on its ability to create value from its assets. Thereby, this is a very relevant ratio of bankruptcy risk.

Market value of equity/Book value of debt: Equity is measured by adding the market value of all stocks, while debt includes both current and long-term liabilities. This is a measure of solvency by showing how much a firm´s equity can decline before the liabilities exceed the value of the assets. In addition, this ratio adds a market value dimension that many former studies did not include.

Sales/Total assets: This is the firm´s the capital turnover ratio. It measures the sales generating ability of the firm´s assets. The ratio is unimportant for this study as it was later dropped by Altman when adapting the model to non-manufacturers.

Altman performed an F-test to test the individual discriminating ability of the ratios. The F-statistic presented in Table 10 shows that ratios X_1 to X_4 are significant at a 1 percent level. Ratio X_5 does not show a significant difference between the groups.

Table 10: Descriptive statistics Z-model from Altman (1968)

Variable	Bankrupt Group Mean	Non-Bankrupt Group Mean	F Ratio	Scaled Vector	Ranking
X_1	-0.061	0.414	32.60*	3.29	5
X_2	-0.626	0.355	58.86*	6.04	4
X_3	-0.318	0.153	26.56*	9.89	1
X_4	0.401	2.477	33.26*	7.42	3
X_5	1.500	1.900	2.84	8.41	2

(*) indicates significant at a 1% significance level

The scaled vector shows the relative contribution of each ratio to the total discriminating power of the function. This way we are able to evaluate each ratio´s contribution on a relative basis. It is computed by calculating each variable´s coefficient with its standard deviation. Table 10 shows that ratio X_3 contributes the most to group separation, followed by X_5, despite its non-significant F-ratio.

For any predictive model, there will be classification errors. As mentioned earlier, there are two types of errors, and it is important to separate them because they might have different costs. Generally, a Type 1 error is considered more serious because the costs related are expected to be higher.

The results of Altman´s Z-score model were tested several ways (Altman E. , 1968). Using the original sample one year prior to bankruptcy the model was extremely accurate, classifying 95% of the firms correctly. The Type 1 error was 6%, while the Type 2 error was 3%. However, because the original sample was used, this model should be considered describing rather than predictive.

Next, the model was applied to data two years prior to bankruptcy. The overall accuracy rate fell to 83%. The Type 1 error rate was 28% while the Type 2 error rate was only 6%. The model was also tested with data up to five years prior to bankruptcy. The accuracy fell significantly for every year added, categorizing only 36% correctly five years prior to bankruptcy.

In order to fully test the predictive ability of the model, secondary samples were collected. First, the model was tested with 25 bankrupt manufacturing firms with similar asset-size range as the initial sample. The model predicted 24 firms correct, resulting in a 96% hit rate. Additionally, a secondary sample of 66 non-bankrupt firms were collected. They were all manufacturing firms that had experienced negative profits during the last three years. The accuracy rate with this sample was 79%.

In order to make the model more applicable to its users Altman defined cut-off points for the Z-score. Firms with a Z-score higher than 2.99 are in the non-bankrupt category, while firms with a Z-score lower than 1.81 are in the bankrupt category. Firms with a score in-between are put in the grey zone, implying that the model cannot conclude on its prediction.

4.1 Adapting the model for private firms

The original Z-score model was limited to listed firms, due to the inclusion of market value of equity (Altman E. , 2000). In order to include private firms, the market value was replaced by book value of equity in ratio X_4. The remaining ratios were not changed. Subsequently, the coefficients were re-estimated, yielding the following model:

$$Z' = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.420X_4 + 0.988X_5$$

For the revised models, Altman calculated the ratios using decimal form, while using percentage form in the original model. Because of this, the coefficients are higher in the revised models.

Table 11: Descriptive statistics Z´-model from Altman (1983)

Variable	Mean Bankrupt	Mean Non-Bankrupt	Univariate F	Scaled Vector	Ranking
X_1	-0.061	0.414	32.6*	0.067	5
X_2	-0.626	0.353	58.8*	0.121	4
X_3	-0.318	0.153	26.6*	0.318	1
X_4	0.494	2.684	25.8*	0.203	3
X_5	1.503	1.939	2.8	0.291	2

(*) indicates significant at a 1% significance level

Compared to original model, the new coefficients have changed. The coefficient of the revised ratio, X_4, decreased from 0.600 to 0.420, when adjusting for the difference in decimal and percentage form. However, the overall properties of the model are fairly similar. The F-ratio show that ratios X_1 to X_4 are significant on a 1 percent level, while X_5 is not significant. The scaled vectors show the same order of contribution to the models discriminating ability.

The model was only tested with the original sample. The results from the revised model were slightly less accurate than the original model. The overall accuracy rate was 94%, with the Type 1 error rate at 9% and the Type 2 error rate at 3%. There was no out-of-sample test due to the lack of data for private firms.

4.2 Adapting the model for more industries

The two previous models were both intended for manufacturing firms. The next step in the development of the Z-score model was to adapt it for non-manufacturers (Altman E. , 2000). The ratio X_5, asset turnover, was identified as a particularly industry-sensitive ratio. As a result, this ratio was dropped, yielding a model with only four ratios. The book value of equity, introduced in the first revision, was also included in this model. After a re-estimation of the coefficients, a new model was presented:

$$Z'' = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$$

Table 12: Descriptive statistics Z´´-model from Altman (1983)					
Variable	**Mean Bankrupt**	**Mean Non-Bankrupt**	**Univariate F**	**Scaled Vector**	**Ranking**
X_1	-0.061	0.414	32.6*	0.267	2
X_2	-0.626	0.353	58.8*	0.205	4
X_3	-0.318	0.153	26.6*	0.304	1
X_4	0.494	2.684	25.8*	0.224	3

(*) indicates significant at a 1% significance level

The F-ratio on a univariate level shows that all the ratios are significant on a 1 percent level. The only ratio not significant in earlier models was now dropped. The scaled vectors show that the ratios contribute a quite similar amount to the overall discrimination. X_1 were now ranked as the second most contributing ratio, as opposed to the least contributing ratio in the earlier models.

When applied on the original sample, the model produced the same accuracy as the Z'-model. The overall accuracy was 94%, with Type 1 error rate at 9% and Type 2 error rate at 3%. The cut-off scores are different from the original model. A score lower than 1.10 is defined as distressed, the non-distressed zone comprises scores higher than 2.60, and the grey zone includes scores in-between. Altman did not test this model on a hold-out-sample.

Even though the revised model produced accurate results for non-manufacturers, it is reasonable to challenge the procedure used to adapt the model. Researchers agree that the ratio dropped is sensitive to industry-specific characteristics, and it can be argued that Altman was correct in dropping this ratio. However, the predictive ability of the remaining ratios was not discussed. We argue that the four remaining ratios were carefully selected based on a sample of manufacturing firms. Hence, there is no reason for these ratios to be superior outside their intended scope.

As stated previously, the Z´´-model is the model we consider when performing our analysis. This is because our data includes private firms from multiple industries.

5. Variable selection and discussion

In order to emulate Altman's study, we use the same classes of ratios. Because we include non-manufacturers, we only consider the three classes included in Altman's Z´´-model. These are liquidity, profitability, and solvency.

We have an analytical approach when producing our sample of ratios. Our selection is based on an analysis of whether the ratios considered are appropriate to measure the aspect they aim to evaluate. In total, our sample consist of 14 potential ratios. The reason for not including more ratios is the difficulty of finding appropriate ratios and subsequently their accompanying data. The selection of these ratios is discussed later in this section.

The approach to selecting the final four ratios is similar to that of Altman. We also put considerable emphasis on the accuracy of the different combinations of ratios. However, we also consider other aspects. The selection criteria are discussed in the method section, while the final profile of ratios is discussed in the results section.

Table 13 summarizes the ratios used by Altman, the ratios not used, and our alternative ratios. The ratios are categorized after class. According to Altman's classifications, none of his four final ratios are considered to be a leverage measure (Altman E. , 1983).

In the following sub-sections, we discuss the limitations and problems of the four ratios in the Z´´-model. Additionally, we evaluate whether the alternative ratios are better. The abbreviations used for Altman's ratios and the alternative ratios are presented in Table 14. These are used throughout the rest of the study.

We have found limited literature on the subject of criticizing Altman's ratios with respect to his analysis. There is also limited literature on the subject of using alternative ratios. The discussion is therefore largely based on a deductive approach where we analyze each ratio.

The alternative ratios that we present were not considered by Altman. However, the Cash and Marketable Securities/Current Liabilities ratio is similar to our Financial Assets/Current Liabilities. Altman considered his ratio to be a measure of liquidity while we use our ratio as a measure of solvency. To the best of our knowledge, there has also been limited use of our alternative ratios in other bankruptcy studies.

Table 13: Summary of Altman's original 22 ratios and our 14 alternative ratios

	Liquidity	Solvency	Profitability	Leverage
Z´´	• Working Capital/Total Assets	• Retained Earnings/Total Assets • MVE/BV of Liabilities	• EBIT/Total Assets	-
Not included by Altman	• Current Assets/Current Liabilities • Cash and Marketable Securities/Current Liabilities	• Net Worth/Liabilities	• Sales Profit/Sales • Earnings Before Taxes/Sales • Net Income/Sales • Earnings After Taxes Before Interest/Total Assets • Number of Years of Negative Profits in Last 3 Years	• Current Liabilities/Total Assets • Non-Current Liabilities/Total Assets • Liabilities/Total Assets
Alternatives	• Net CFO minus Net Investments/Interest • EBITDA minus Net Investments/Interest • EBIT/Interest	• Financial Assets/Liabilities • Financial Assets/Current Liabilities • BVE minus Goodwill/BV of Liabilities • Financial Assets/Worst Case Cash Outflow	• Avg. Net Income/Avg. Total Assets • Avg. EBITDA/Avg. Total Assets • Avg. EBIT/Avg. Total Assets	• EBITDA/Invested Capital • EBIT/Invested Capital • EBITDA/Operational Assets • EBIT/Operational Assets

Table 14: Abbreviations for all ratios	
Notation	**Ratio**
WC/TA	$\frac{\text{Working Capital}}{\text{Total Assets}}$
RE/TA	$\frac{\text{Retained Earnings}}{\text{Total Assets}}$
EBIT/TA	$\frac{\text{EBIT}}{\text{Total Assets}}$
BVE/BVL	$\frac{\text{Book Value of Equity}}{\text{Book value of Liabilities}}$
CFO/I	$\frac{\text{Net CFO} - \text{Net Investments Operating Assets}}{\text{Interest Costs}}$
EBITDA/I	$\frac{\text{EBITDA} - \text{Net Investments Operating Assets}}{\text{Interest Costs}}$
EBIT/I	$\frac{\text{EBIT}}{\text{Interest Costs}}$
ANI/ATA	$\frac{\text{Average Net Income}}{\text{Average Total Assets}}$
AEBITDA/AOA	$\frac{\text{Average EBITDA}}{\text{Average Operational Assets}}$
AEBIT/AOA	$\frac{\text{Average EBIT}}{\text{Average Operational Assets}}$
EBITDA/IC	$\frac{\text{EBITDA}}{\text{Invested Capital}}$
EBIT/IC	$\frac{\text{EBIT}}{\text{Invested Capital}}$
EBITDA/OA	$\frac{\text{EBITDA}}{\text{Operational Assets}}$
EBIT/OA	$\frac{\text{EBIT}}{\text{Operational Assets}}$
FA/L	$\frac{\text{Financial Assets}}{\text{Liabilities}}$
FA/CL	$\frac{\text{Financial Assets}}{\text{Current Liabilities}}$
BVEG/BVL	$\frac{\text{Book Value of Equity} - \text{Goodwill}}{\text{Book Value of Liabilities}}$
FA/WCF12	$\frac{\text{Financial Assets}}{\text{Worst Case Cash Outflow Next 12 Months}}$

5.1 Liquidity

Liquidity is necessary to cover ongoing expenses and is generally associated with the ability to meet short-term obligations. A ratio that considers the health of the company from a liquidity perspective should therefore consider whether the company has the funds to cover ongoing expenses. We put particular emphasis on the ratio being appropriate when assuming continued operations.

5.1.1 Criticism of Working Capital/Total Assets

According to the meta study by Gissel et al. (2007), this ratio is the second most popular liquidity ratio across 165 bankruptcy studies. It is slightly less popular than the current ratio and the third most popular ratio overall. According to Altman (1983), it is the second most important factor in the Z´´-score model.

The ratio expresses the relationship between the working capital and the total assets of the firm. An increase in current assets would increase the ratio. Similarly, a reduction in current liabilities would also increase the ratio. The current liabilities represent a source of default risk to the firm. If the current assets are able to cover current liabilities, the short-term risk of default is reduced. However, this relationship does not account for the nature of the items included in current assets and current liabilities.

First of all, inventory is an important item in current assets. The inventory is essential for most firms in order to operate. Therefore, it makes little sense to sell the inventory to cover current liabilities, as the firm will have to take up new debt in order to finance new inventory.

Accounts receivable face a similar problem. When customers have the opportunity to postpone payments, accounts receivable increases. This means that accounts receivable will continue to increase as the company collects outstanding payments in order to cover liabilities. Both of these items often make up a significant part of the current assets, but are essentially useless in order to cover current liabilities as long as operations continue.

Several items in current liabilities face similar problems, when we assume continued operations. Items such as accounts payable, taxes, provisions on wages etc. are continually being repaid and accrued. The nature of many current liabilities is regenerative as long as operations continue.

Another issue is the estimates of current assets and current liabilities. A firm in a situation that requires the sale of current assets does not have a good bargaining position. It is reasonable to assume that a price negotiated under these conditions is below the book value. This problem is exacerbated further by the fact that some current assets might be tailored specifically to the firm, and are thereby less valuable to others. Therefore, it makes more sense to use liquidation values for current assets. Similarly, if the firm is defaulting, all liabilities become current. Only using the current liabilities produces a value that is too small. From this, we see that the relationship between current assets and current liabilities can be inaccurate.

Since working capital is an absolute value, it is necessary to adjust for the relative size of the firm. This ratio uses total assets to adjust for size. The implication of this is that an increase in working capital, all things being equal, will reduce the default risk. This is consistent with the discussion above. However, this also indicates that a reduction in total assets would, reduce the risk of default. This is less intuitive and the limitations and problems of this relationship are discussed below.

A limitation of using total assets as the denominator is that it does not take into account the reason behind a change in total assets. For instance, it is generally accepted that a high equity to liability ratio decreases the risk of default. Hence, if equity decreases, for example as a result of an amortization of a non-current asset, the risk of default should increase. However, a decrease in equity would, all things being equal, result in a decrease in total assets. According to the ratio, this should reduce the risk of default. As we can see, this interpretation is counterintuitive.

Additionally, the denominator indicates that smaller asset firms, are less likely to default. This is not necessarily the case as the book value of assets varies significantly between industries. Some industries are capital intensive and have relatively high book values. Other industries are more labor intensive, where most of the firm value is outside the balance sheet, resulting in a relatively low book value of total assets. The rate of default varies greatly between these industries and it is difficult to draw the conclusion that smaller firms are less likely to default.

Lastly, it is important to notice that the ratio dictates that the ideal composition between current and non-current assets is a large share of current assets. This is not accurate as the ideal compositions varies depending on the industry. Although the ratio is not supposed to consider this relationship directly, it is an indirect effect which comes from including current assets in the nominator and total assets in the denominator.

From this discussion, it is apparent that WC/TA might not be the best way to evaluate a company's financial health with respect to liquidity. Hence, we suggest using an alternative ratio.

5.1.2 Alternatives to Working Capital/Total Assets

<u>Net CFO minus net investments in operations/Interests and instalments</u>

Using net cash flow from operations (CFO) for evaluating liquidity is appropriate as the cash generated throughout the year is what the company can use to cover its expenses, without borrowing from lenders. Contrary, current assets are often essential for ongoing operations, and are therefore indispensable. Hence, it is more realistic to use net CFO as a source to cover the company's expenses.

An increase in net CFO directly increases the funds available to cover expenses. Therefore, companies with relatively high net CFO have more liquidity and the risk of default is smaller, all other things being equal. This is in contrast to using current assets, where an increase in current assets does not necessarily increase funds available to cover expenses. Increases in inventory and accounts receivable are examples of this, as discussed earlier.

As we work under the assumption that operations continue, we need to account for net investments in operations. This is because a company needs to reinvest as they grow, assets deteriorate and technologies change. If a company does not reinvest, operations will eventually cease as a result of deterioration or competition. This will result in the company defaulting. Not including investments gives an unrealistic and inflated measure of how well the company can cover its expenses.

Net CFO and investments are absolute measures, and it is necessary to evaluate them relative to a suitable measure. One could measure the available funds to cover expenses relative to the expenses they are supposed to cover. Therefore, we suggest measuring net CFO minus investments relative to interest and instalments. Because the nominator includes cash flow items, it is preferable if we use interest payments and instalment payments, not costs. Using interest and instalments ensures that the focus remains on the actual expenses the company has to cover. Similarly, using interests and instalments is more in line with what the ratio tries to measure, i.e. default risk with respect to liquidity. A well-run company usually has the possibility to refinance instalments through new debt. Hence, it is possible to focus on interest coverage alone.

As interests and instalments increase, the ratio decreases, indicating that the risk of default is higher. This makes intuitive sense as higher interest and instalment directly increase the risk of the company not being able to repay its lenders, and thereby defaulting on its debts. Additionally, high interest and instalments expenses indirectly indicate that the company has relatively high liabilities, which in turn increases the risk of default. Hence, we argue that using interest and instalments is more appropriate than using total assets.

One major weakness with regards to this ratio is that cash flows usually exhibit large fluctuations between years. The source of these fluctuations is often outside the company's sphere of influence, such as when payments are received or when payments need to be made. Since we use one year's performance to evaluate the risk of default, we risk not giving an accurate depiction of the financial reality of the firm. This limitation is prevalent in all one-year prior models, but the effect is exacerbated by the volatile nature of cash flows.

Furthermore, if we ignore instalments we are dependent on the company being well-run. This represents a limitation in the ratio, as this might not be the case if the company is facing bankruptcy. There is a risk that the ratio does not effectively assess the liquidity state if a company is close to bankruptcy, and the accuracy of the model might suffer from this.

<u>EBITDA minus net investments in operations/Interest and instalments</u>

Another alternative is using EBITDA, instead of net CFO, to represent the funds generated throughout the year. EBITDA, and similar accounting measures, do not represent the actual cash generated by the company. Hence, EBITDA is not a direct representation of the funds available to cover expenses. Despite that, earnings should over time be converted to cash and could therefore serve as a proxy for net CFO. As long as there is a sufficient correlation between earnings and cash flows, it is reasonable to use EBITDA to represent generated cash.

An increase in EBITDA would decrease the risk of default with regards to liquidity, as long as earnings are converted to cash. If this assumption does not hold, the estimated default risk may be inaccurate.

As stated earlier, we assume that operations continue and investments are needed as a result of this. We use net investments in operations to represent these investments, similarly to the previous alternative ratio. However, we note that EBITDA is an earnings measure, while investments in operations is a cash flow measure. It might be more prudent to represent expenses with regards to investments using accounting measures. One possibility is using depreciation and amortization as a proxy for investments, leaving us with EBIT as the nominator. This will ensure that there is a consistency between the items in the ratio.

An increase in net investments in operations would reduce the value of the nominator, hence increasing the risk of default. This makes intuitive sense as less funds will be available to cover expenses. Similarly, an increase in depreciation and amortization would also increase the default risk. The predictive accuracy of this change is however dependent on whether depreciation and amortization are good proxies for net investments.

Again, we suggest using interests and instalments in the denominator. Additionally, we stress the fact that a well-run company has the option to refinance instalments, and hence it is possible to look at interest coverage alone. Because we use an accounting measure, it is more prudent to use interest and instalment costs as opposed to payments. However, this creates similar problems to those of using EBITDA instead of net CFO, namely that costs do not necessarily represent actual expenses. Nevertheless, as long as there is a correlation between costs and cash outflows following these items, the predictive ability should remain relatively similar.

As touched on above, using accounting measures instead of cash flows represents a potential source of error, as they do not necessarily give an accurate description of the financial situation. Lenders are concerned with the actual interest and instalment payments, not how the firm decides to expense these items. Similarly, only the actual cash generated can be used to cover these payments. Therefore, there is a source of error in using accounting measures. This is exacerbated if the correlation between accounting measures and cash flows is small.

However, using accounting measures may remove a major weakness of cash flow measures, namely the volatile nature of cash flows. Using EBITDA could for instance represent a “normalized” net CFO. The same is the case for using interest cost as opposed to interest payments. This could reduce the effect random factors outside the companies sphere of influence have on the predictive ability of the ratio. Hence, there is an argument for using accounting measures, especially since we use a one-year prior model.

5.2 Solvency

Altman classified two ratios as solvency measures: RE/TA and BVE/BVL. Solvency considers whether a firm is able to meet its long-term financial obligations. A firm is generally considered solvent as long as the value of the assets is greater than the value of the liabilities. The two ratios measure different aspects of solvency, and are as such discussed separately. Alternative ratios are also discussed separately as these too measure different aspects.

5.2.1 Criticism of Retained Earning/Total Assets

Gissel et al. (2007) found that this ratio was the fourth most popular ratio overall, across the 165 studies they examined. Consequently, it is the most popular ratio for measuring historical profitability and age. The ratio was the least important ratio in the original Z´´-model.

The ratio evaluates solvency from different angles. First of all, it considers the size of equity relative to total assets, as retained earnings are included in equity. This is similar to other leverage ratios that measure solvency. It also considers solvency with respect to historical profitability, where a greater historical profitability indicates a more solvent firm. This is similar to interest coverage measures that evaluate solvency; greater retained earnings indicate that the company has been able to cover interests over time.

Considering the popularity of the ratio it arguably has some informational value. We discuss the logic behind using the accounting figures included in the ratio, as well as the problems and limitations with the figures.

The ratio considers the size of the retained earnings relative to the company's size. Retained earnings is defined as the sum of net income after dividends payments and extraordinary items, over time. An increase in retained earnings improves the ratio, indicating a decrease in default risk.

In order for retained earnings to increase, the company needs to be profitable over time. Historic profitability indicates that the company is well run and that it may have some competitive advantages. These types of companies are arguably less likely to default than companies that historically have been unprofitable. In this case, the ratio is intuitive and coherent with realistic expectations for default risk.

One issue with retained earnings is that they are subject to dividend policy. Companies may have different retained earnings even though the underlying economical features may be the same. This essentially means that retained earnings represent historical profitability adjusted for dividend policy. Some companies may be less profitable, but given their conservative dividend policies the ratio may categorize them as more historically profitable than what they are. Therefore, one should consider retained earnings adjusted for historical dividends in order to accurately measure historical profitability.

Furthermore, companies may restructure their balance sheets. One possible result is the removal of a company's deficit in retained earnings. Another result may be the restructuring of retained earnings as equity. Hence, some companies may have retained earnings that do not accurately express past results. This is especially problematic as a restructuring of the balance may indicate that the financial situation of the company is far from optimal.

Another issue with using retained earnings is whether historical profitability is a good indicator of future profitability. Relatively old companies, that have been historically profitable, might have a higher risk of bankruptcy if they for instance are highly invested in an unprofitable industry. The ratio fails to consider this, and it may over-evaluate these types of companies.

With regards to the denominator, Altman again uses total assets. The issues with the denominator are similar to those discussed for the previous ratio. It implicitly indicates that smaller firms are less likely to default, as a decrease in Total Assets increases the ratio. Additionally, it may not be an accurate estimator of size in today's business environment where more companies have relatively small balance sheets. It also fails to consider what causes the change in total assets.

However, it is arguably more fitting to use total assets as the denominator in this case than for the other ratios. This is because RE/TA measures how much of the total assets that are financed through retained earnings. Hence, there is consistency between the nominator and denominator. This is not the case for the other ratios.

Given these issues, we argue that other ratios might be better suited to assessing solvency from the perspective of historical profitability. We will evaluate the alternatives below.

5.2.2 Alternatives to Retained Earnings/Total Assets

<u>Average Net Income/Average Total Assets</u>

This alternative ratio uses the three-year moving average of net income and total assets to assess the recent historical profitability of a firm. The choice of a three-year average is a result of our discretion and the data available. An average including a different number of years is also possible. A high ratio indicates that the company has been profitable in recent years and thus would have a decreased risk of defaulting.

The benefit of this ratio is that it removes the noise caused by dividends policies and restructuring of the balance sheet. It also limits the time horizon, hence removing any noise caused by results that are no longer representative. The ratio is consistent as it measures the net income relative to total assets. Using total assets to adjust for the size of the company is more appropriate in this case, as net income is the sum of incomes and expenses generated by all assets.

One weakness is that some information may be lost when using a limited time period. Furthermore, the ratio is again dependent on historical data being an indicator of the future. Net income may contain items that are non-repetitive and the value might be distorted by such items. A struggling company may for example sell assets, resulting in an unusually high net income in one year.

Average EBITDA/Average Operational Assets

Similarly to the discussion above, this ratio uses a three-year moving average to assess the recent returns of a firm. A high ratio indicates that the company has had strong return on their core business and thus should have a decreased risk of defaulting.

The results are measured relative to operational assets, which are defined as the assets that are used in the firm's core business. If operational assets increase without a subsequent increase in EBITDA, the firm is using its assets less efficiently. This indicates a higher risk of return and is consistent with realistic expectations for bankruptcy risk.

The ratio is very similar to using ANI/ATA and shares many of the same strengths and weaknesses. However, one major difference is that using EBITDA removes the noise caused by non-repetitive items. Additionally, EBITDA can be substituted with EBIT in order to include depreciation and amortization.

One weakness of this ratio, compared to the pervious alternative, is that it does not include all items. Increases in bankruptcy risk can also be triggered by developments outside the firm's core operations. This indicates that valuable information can be lost when using EBIT or EBITDA, compared to net income.

5.2.3 Criticism of Book Value of Equity/Book Value of Liabilities

This ratio is one of the least used in the study by Gissel et al. (2007). This is largely because the use of market values for equity is preferred. The ratio was the second least important in the Z´´-model.

This ratio evaluates different aspects than the other solvency ratio, RE/TA. They both evaluate solvency with regards to leverage. However, this ratio considers the entire equity. Furthermore, this ratio considers its financial obligations by including liabilities. This is in contrast to RE/TA, which considers historical profitability and indirectly historical coverage ratio. In general, the BVE/BVL ratio considers the robustness of a company to a greater degree than RE/TA.

Given that this specific measure is not frequently used, we could assume that the predictive information it offers is limited. On the other hand, the market value of equity is quite popular with regards to bankruptcy studies. This indicates that there is a belief that the equity/liabilities relationship offers valuable information with regards to default risk. Below we discuss the general idea behind using the ratio. We will also highlight some of the limitations and problems behind this ratio.

The equity to liabilities ratio is classified as a solvency measure. It evaluates the robustness of a firm and whether the firm is able to meet its financial obligations. The explicit interpretation of the ratio is to what degree the company can cover its liabilities through its equity, and to what degree the company is leveraged. Implicitly, it follows that a high degree of equity indicates a low risk of default.

Generally, highly leveraged companies have a higher default risk. Additionally, a company able to cover all its liabilities using equity is generally better suited to handle a situation where several liabilities fall due in a short period of time. Although this ratio intuitively seems appropriate to measure default risk, it fails to consider the nature of the equity.

The most important limitation is that equity is the residual value of the total assets, when liabilities have been subtracted. Consequently, using equity to cover liabilities means selling assets. Another limitation is that the value of assets, and thereby equity, can be measured using different methods.

As mentioned when discussing WC/TA, we need to stress the problem of selling assets without impacting operations. Reinvestments in assets such as inventory, equipment, plants etc. are necessary to continue operations. Unless an influx of capital is possible, the sale of an operating asset to cover liabilities will require the company to take up debt to repurchase a similar asset. This indicates that using equity to cover liabilities is problematic as some assets are necessary in order to operate. It also indicates that equity is overvalued if it is considered to be a proxy for whether the company can cover its liabilities.

Some assets are simple to value accurately at fair value, such as cash and financial instruments. Others are more challenging, and many different approaches are available, such as estimated fair value, historical cost, amortized cost, realizable value, and expected value. A single item can often be evaluated using more than one approach. This represents a potential source of error when valuing assets and comparing firms.

Similarly, it is important to consider that the values the firm can expect to obtain might be different if the firm is under pressure. Some assets should be valued closer to liquidation value in order to represent a realistic value. This is also the case if some assets are highly specialized to the firm’s needs. It is an important aspect to consider, as selling assets often is a viable option to cover expenses in order to avoid bankruptcy.

Additionally, a limitation is related to which accounting items that make up the equity. We need to consider the nature of the assets in order to evaluate whether the values of the assets represent a realistic source of debt coverage. If the items do not represent a realistic source of financing, they should be excluded. If they are not excluded they will overvalue the financing available to cover liabilities.

As discussed above, some items are problematic to sell because they make up the daily operations of the firm, whereas intangibles are problematic by their very nature. Goodwill can make up a significant amount of the total assets of a firm, but it is practically worthless as a mean to cover liabilities. Patents and research development face a similar issue. They may be worthless to other firms unless a significant amount is sold off, which would often severely impact daily operations. Ongoing contracts also face a similar issue, as customers may have objections or be entitled to compensations. All these factors indicate that using equity as a proxy for a "rainy day fund" would give an inflated estimate of how well the company can cover its liabilities.

From the discussion, we see that the equity to liabilities ratio has some limitations that make it a problematic predictor of default risk. Often equity is an inflated measure of how well the company could cover its liabilities, unless the equity in question is treasury shares. Its redeeming factor is that it gives an indication of the degree of leverage. We recommend using a ratio that considers the true driver of the default risk, namely whether the company has assets that can cover its liabilities.

5.2.4 Alternatives to Book Value of Equity/Book Value of Liabilities

Financial assets and unused credit/Worst case net cash outflow next 12 months

Financial assets are, by definition, assets that can be sold without affecting operations. Therefore, they represent a source for the firm to cover potential losses. Hence, it is reasonable to use financial assets to estimate the robustness of the firm. Additionally, financial assets are often measured accurately to real market values. This means that they also represent an accurate estimate of robustness. Similarly, any lines of unused credit represent a source of financing that can be used to cover losses if necessary. On the contrary, the book value of equity may consist of assets that are intangible, associated with operations or difficult to measure accurately to market values. Therefore, financial assets are better suited to measure robustness.

When the value of financial assets increases, the value of the assets that the firm can sell to cover losses, without a negative impact, directly increases. The same is the case for unused lines of credit. This is in contrast to an increase in the book value of equity, which might be driven by an increase in goodwill or a revaluation of an operating item. These are items that are difficult to convert to cash or are important for continued operations.

Furthermore, using the worst-case net cash outflow as the denominator directly represents the loss that needs to be covered. From the ratio, we see that an increase in the worst-case scenario indicates a higher risk of default. This is reasonable as larger losses are more difficult to absorb.

Using worst case net cash outflow instead of liabilities is not necessarily better. We also discuss measures that use liabilities as the denominator for measuring robustness later. However, one should note that using liabilities as the denominator might be inaccurate to some extent. All liabilities are rarely due at once, unless the company is facing bankruptcy, at which time it might too late to perform such analysis. Therefore, they don't necessarily represent the losses the firm has to absorb during bad times. This might affect the predictive power of the ratio.

A significant problem with this ratio is estimating unused lines of credit and worst-case net cash outflows. Unused lines of credit might be found in the financial statement and could therefore be relatively simple to estimate. However, if the financial statements don't contain this information, it might be difficult to estimate. Worst case net cash outflows are difficult to estimate as they and require in-debt analysis. Hence, they are time consuming and vulnerable to discretions made by the analyst. Using historical values is possible, but these can be inaccurate as there is no guarantee that future losses are similar to historical losses. Although this ratio represents an accurate estimate of robustness in theory, it has some limitations when applied practically.

Financial assets and unused credit/Liabilities

Again, we use financial assets and unused credit as the nominator to estimate the robustness of the firm. However, we now use liabilities as the denominator. This is the same denominator that Altman used for his ratio. Although liabilities do not directly represent the losses a firm has to absorb, they offer some information that can be beneficial when estimating robustness.

An increase in liabilities usually represents an increased risk of default. Hence, there is an intuitive interpretation of the change in liabilities with respect to default risk. Furthermore, if the loss the firm experiences is significant, the firm might not be able to cover its obligations to the lenders. This results in a technical default, where all liabilities fall due within a year. If this is true, it would be prudent to evaluate the firm's ability to manage such a situation. In this situation, an increase in liabilities would represent a higher risk of default.

It is possible to use current liabilities instead of total liabilities as the denominator. The arguments for using current liabilities are similar to those of total liabilities. An increase in current liabilities represents an increased risk of default. Similarly, if the loss is significant, the firm might struggle to uphold its obligations. Evaluating to what degree financial assets can be used to cover current liabilities in this situation, represents a measure of the robustness of the firm.

Although using liabilities instead of worst case net cash outflow might be less accurate, it might be necessary in practice. Liabilities and current liabilities are far easier to estimate than the worst-case net cash outflow. They also contain some additional information that the net cash outflow does not. This makes them a viable alternative to net cash outflow.

5.3 Profitability

It is essential for a firm to be competitive with respect to its return. Firms that produce returns that are lower than those of comparable firms have a weaker income to cost relationship and/or are less efficient. These firms naturally have a higher risk of default. When assessing the ratios, we put emphasis on core operations, as these represent the long-term results. Additionally, we focus on there being consistency between the nominator and denominator. This is because we want to measure income and expenses using the assets that are responsible for generating these.

5.3.1 Criticism of EBIT/TA

Gissel et al. (2007) found this ratio to be fifth most used ratio across all 165 studies. It is the second most popular for measuring return on investment. This ratio has the largest scaled vector in the Z´´-model, and is hence the most important.

The relative importance of the ratio indicates that it possibly contains important information with regards to the default risk of a firm. This is arguably the case as the ratio expresses an important aspect of a firm's financial health; whether it is profitable and competitive with respect to its operating activities.

An increase in EBIT would, all things being equal, increase the ratio and hence reduce the risk of default. This is reasonable as it indicates an increase in income or decrease in costs. This type of change indicates an improvement in the income to cost relationship and is expected to reduce the risk of default. If the change is driven by an increase in income it also represents an improvement in productivity, as the firm is able to generate more income from a given pool of assets. This is also expected to reduce default risk. Hence, there is an intuitive and logical relationship between a change in the nominator and the real default risk.

EBIT mainly contain operating items. These items usually have a high predictive value, based on the fact that they are repetitive by nature. However, EBIT also contains non-operating items, mostly classified under non-operating income and costs. These items are arguably less repetitive in nature. Hence, including these might reduce the predictive ability of the ratio.

A study by Cheng, Cheung and Gopalakrishnan (1993) evaluated how well operating income, net income and comprehensive income explains residual security returns. They found that operating income has more information content than net income, while the latter has more information content than comprehensive income. Based on this, and other similar studies, it is possible to extrapolate that non-operating items have an adverse effect on the information content.

It follows from the discussion above that the inclusion of non-operating items in EBIT could have an adverse effect on the accuracy of the ratio. However, if these items appear to be predictable and repetitive, the possible adverse effect may be reduced. Another possibility is to normalize EBIT with respect to these special items, in order to produce a more representative measure of the firm's results.

We have previously stressed the need to focus on continued operations when assessing ratios. Hence, when measuring return on investments, it is important to include any costs associated with reinvestments. By using earnings after depreciation and amortization Altman has, to some degree, accounted for this. These items can serve as a proxy for reinvestments, as the company's reinvestments are ideally at least equal to depreciation and amortization. Additionally, amortization has a signaling effect, as amortizations may indicate financial distress.

The problem with including depreciation and amortization is that these items are subject to the management's discretion. First of all, the depreciation method can vary between different firms. Second, the depreciation time can also differ, even for similar items. Lastly, amortization can differ substantially between firms, especially with respect to how they value amortized items. All of this represents a substantial source of error in the model.

As a worst-case scenario, including depreciation and amortization can leave the model susceptible to errors from earnings management. Gerety and Lehn (1997) argue that accounting fraud is more prevalent when it is costly to verify the quality of a good and that fraud is correlated with the degree of asymmetrical information. Both of these characteristics can be associated with depreciation and amortization.

The return is measured relative to the total assets of the firm. The implication of this is that an increase in total assets reduces the ratio, which indicates an increase in default risk. This is intuitive with regards to efficiency. If total assets increase, without a related increase in income, the company is using its assets in a less efficient manner. However, an increase in total assets indicates that the size of the firm has increased. This implicates that larger firms have a higher risk of default than smaller. This relationship has been discussed earlier and it is arguably less intuitive.

One problem with using total assets in the denominator is that all assets, including financial assets are included. This is problematic as EBIT does not include financial income or costs. Hence, there is a lack of consistency between the nominator and denominator, which might affect the predictive ability of the ratio.

Furthermore, financial assets could in theory be sold with the purpose of reducing liabilities without effecting operations, thereby reducing the size of the balance sheet. This indicates that firms where financial assets make up a relatively large portion of total assets are less efficient. This is, as we can see, not necessarily the case. Hence, the ratio might not accurately predict the true productivity of the firm

Considering these issues, we suggest using a different ratio for determining the return on investment. We focus on reducing the potential source of error from the included items and increasing the consistency between the measures we use.

5.3.2 Alternatives to EBIT/Total Assets

<u>EBITDA/Invested Capital</u>

This ratio uses EBITDA to represent the result of the firm. It consists of income and costs from operations, which are proven to have a high predictive ability. An increase in EBITDA results in a higher ratio and thereby a lower risk of default. The increase can be driven by an increase in income, a decrease in costs, or both. This naturally improves the firm’s financial situation. Hence, there is consistency between how the change affects the ratio and the effect the change has on the real default risk. EBIT is also a possible measure to be used as the nominator. As discussed previously, there is an intuitive and logical relationship between a change in EBIT and the real default risk.

A benefit of using EBITDA compared to EBIT is that we eliminate the source of error related to depreciation and amortization discussed earlier. Additionally, this reduces the risk of earnings management affecting the accuracy of the model. It is also possible that using EBITDA can improve the predictive value. This is because we only consider operational income and cost, which should have more informational value than extraordinary items. However, it is not necessarily the case that using EBITDA instead of EBIT will have an effect on the predictive ability of the model. If the issues mentioned above are not significant, the result should be fairly similar.

This ratio uses invested capital to represent the assets used to generate the result. The invested capital is defined as operational assets minus operational liabilities. An increase in invested capital is hence driven by an increase in operational assets or a decrease in operational liabilities. If invested capital increases while EBITDA remains constant, the risk of default increases. This seems logical as it indicates that the company is using its assets in a less efficient matter.

A decrease in operational liabilities results in a lower ratio and an increased risk of default. This relationship is less intuitive than that of a change in operational assets. A relatively small value indicates that a large portion of the operational assets are financed by the firm. Hence, the firm requires a higher return than if the investment in operational assets was highly leveraged. However, an unfortunate result of including operational liabilities is that the ratio implicitly states that a higher degree of leverage reduces the risk of default. This is arguably not the case and could reduce the predictive power of the model. This indicates that using invested capital is appropriate when measuring returns, but not necessarily when assessing the risk of default.

Finally, using invested capital removes the problem of including financial assets that in theory can be sold without affecting the daily operations. It also ensures that there is coherence between the nominator and denominator, with respect to evaluating operational results with operational assets.

EBITDA/Operational assets

Using EBITDA and operational assets have the same benefits as discussed above. Again, it is possible to replace EBITDA with EBIT. This ratio is very similar to using EBITDA to invested capital ratio, with one difference; operational liabilities are removed from the denominator.

When operational assets increase, the risk of default increases. Again, this is because it indicates that the company is becoming less efficient. Using operational assets with EBITDA or EBIT ensures that there is consistency between the nominator and denominator.

As opposed to using total assets we now evaluate operational results using operational assets and we remove the problems related to financial assets.

As mentioned above, including operational liabilities implicitly indicates that being more leveraged reduces the risk of default. This represent a weakness in using EBITDA/IC which we correct by removing operational liabilities. The benefit of this is relatively uncertain.

6. Data

The main sample for this study consists of a balanced sample of 158 Norwegian firms. The original database for the study was provided by the Centre for Applied Research at NHH (SNF). The database consists of financial data from both independent firms and concerns, in the period 1992-2014. The majority of the firms are privately held, but there are some listed firms. The modifications to the data in this study are performed using Stata® 15.

Following the inclusion of both public and private firms, there are differences in the accounting standards among the firms. The EU passed a resolution that required listed companies to prepare financial statements pursuant to the IFRS (Berner, Mjøs, & Olving, 2016). Private Norwegian companies can freely choose between IFRS and Norwegian GAAP.

Differences in accounting standards and valuation might affect our study. Berner and Olving studied how the financial key figures were affected by the differences among IFRS and the Norwegian reporting standard NGAAP (Berner & Olving, 2013). They found some inconsistent and conflicting results, and concluded that there is no substantial evidence that the key figures are affected by IFRS. We are thereby assuming that this does not create a bias for our study.

Ideally, we would want to test two samples, one using IFRS and one using Norwegian GAAP, and see if there were any significant differences in prediction quality. However, the database was inconsistent in reporting whether IFRS was used or not. Therefore, we were not able to perform such an analysis.

The data has been gathered from financial statements submitted to the Brønnøysund Register Centre (Berner, Mjøs, & Olving, 2016). This was carried out by Bisnode D&B in collaboration with Menon Business Economics AS. Thereafter, the files have been standardized and quality assured by SNF in order to create a complete database.

The database coverage has increased over time. In 1992, it included 88 025 companies and 5 891 concerns, while in 2014 it included 289 455 companies and 4 197 concerns. In total, the data base includes 491 783 unique organizational numbers.

If a firm was part of a concern, we used the data for the concern. Similarly, if a subsidiary files for bankruptcy, they are not included. A default for a concern is defined, in this study, when the parent company filed for bankruptcy. This distinction is to ensure that all incomes and costs are included. This is because transactions between companies within a concern might have distinctive effects on the accounting figures.

6.1 Sample selection of bankrupt firms

The first modification we performed to the data set was to adjust all the accounting values for inflation. We chose 2007 as the base year. Hence, all values in our database throughout the study, are presented in 2007 NOK. We found few studies that mention this modification to the data. However, this does not necessarily indicate that they did not adjust for inflation, as it could be the case that they simply do not mention it.

Additionally, we removed firms with missing values for sales and total assets. Missing values for these items were a good indication of poor data quality. Ideally one would consider whether dropping firms with missing values affects the representativeness of the sample, as pointed out by Zmijewski (1984). We did not evaluate whether dropping the firms had an effect on the sample. This is because the majority of the firms that were dropped based on missing values were small. These would have been dropped later, so the effect on the study is arguably insignificant.

In order to conduct our analysis, we need to construct a representative sample similarly to what Altman did. The first restriction we applied was to only include data from 2007 to 2016. This procedure dropped 109 826 firms, and we were left with 381 957 firms.

Altman comments that using a sample of 20 years is sub-optimal (Altman E. , 2000). However, he found it necessary because of limited data. This is the case for our sample as well, as limited data made it necessary to use data spanning several years. It is important to note that the time period we have considered is relatively different from the time period that Altman used. The economic growth was higher and the variation was larger in the U.S. between 1946 and 1965, compared to Norway between 2007 and 2016 (Bureau of Economic Analysis, 2017; The World Bank, 2017).

Next, we removed all non-bankrupt firms. In earlier research, different definitions of bankruptcy have been used. Our database included bankruptcy information gathered from the Register of Bankruptcies at Brønnøysund Register Centre. Thereby, we were applying the official definition of bankruptcy in Norway. Altman also used the official regional definition of a bankruptcy in his study. In both cases, these are defined as when a company files for bankruptcy. After this screening, we were left with 25 835 bankrupt firms.

We then specified the industries we wanted to study. We only dropped firms from the finance and insurance sector due to their characteristics in terms of operation and capital structure. This is fundamentally different from most other industries, and thereby the financial ratios will not be suited to predict bankruptcy. After dropping this sector, we were left with 25 368 firms.

Altman did not specify which industries that were included when he re-estimated the Z´´-model. Therefore, we cannot compare the industries included in our sample with Altman's sample. However, as the model was developed specifically to estimate bankruptcy for non-manufactures it is reasonable to assume that most industries were included. Hence our sample should be relatively similar in composition to his.

Some alternative ratios required cash flow figures. However, this information is not available for all firms. The Norwegian legislation on financial statements, the "Lov om årsregnskap" § 3-2 states that small businesses are exempted from providing cash flow statements. "Lov om årsregnskap" § 1-6 states that a firm is defined as small if it does not exceed two out of three of the following circumstances:

- Sales revenue: NOK 70 million
- Total assets: NOK 35 million
- Average number of employees in the accounting year: 50 full-time year

The database did not include cash flow statements, which made it difficult to filter out the correct firms. We dropped firms based on carefully selected amounts of sales revenue and total assets. The data quality of the number of employees was poor and inconsistent, hence we did not drop on a basis of employees.

After dropping firms with average sales revenue less than 35 million and total assets less than 20 million, we were left with 390 bankrupt firms. Because we used only two out of three specifications, we were not able to fully avoid dropping firms that were not classified as small businesses. However, because we used values that were lower than the values stated by law, it is reasonable to assume that this number is limited.

Altman also had minimum requirements with regards to the size of a firm. The asset size restriction was a minimum of $1 million and maximum of $25 million. If we assume that all the firms included in Altman's sample reported their numbers in 1965 US dollars, the inflation adjusted minimum requirement translates into $6.58 million in 2007 USD. This is equal to NOK 38.56 million in 2007, when using the average exchange rate in 2007. Hence, we see that the minimum requirements are relatively comparable.

We also estimated the average asset size in Altman's sample. This was USD 8 million in 1968, which equals NOK 308.6 million in 2007. In our sample, the average asset size is NOK 144.54 million. The difference between these is relatively large. However, this is expected as Altman used U.S. firms while we use Norwegian firms. Nevertheless, both samples consist of relatively large firms and are therefore comparable.

Because the database did not include cash flow statements, we had to manually search for them. The number of firms with available cash flow information was limited. For this data collection, we used an online data base provided by Forvalt. This database also collects data from the Brønnøysund Register Centre, and we are confident that the data quality of this database is satisfactory (Forvalt).

After adding the available cash flow information, we had our final sample of bankrupt firms. It consisted of 95 firms that went bankrupt between 2008 and 2016. This number was reduced to 79 bankrupt firms due to problems with outliers and missing data. This is greater than Altman's sample of bankrupt firms when he estimated the Z-model. The size of the sample used to re-estimate the Z´´-model is not known.

6.2 Control group

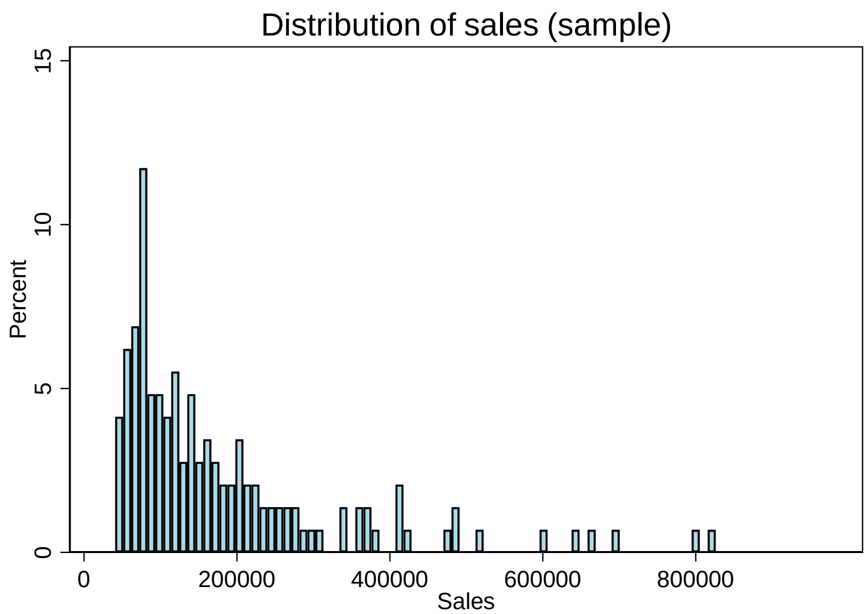
The database provided by SNF was also used to obtain a sample of non-bankrupt firms. The non-bankrupt firms were matched on an industry and asset size basis, similarly to Altman. We found average sales revenue and average total assets for the bankrupt firms in each industry, and randomly selected a group of non-bankrupt firms with similar characteristics. The share of each industry is the same in both the bankrupt and non-bankrupt groups. This way, we are ensuring that the size and industry composition is fairly similar.

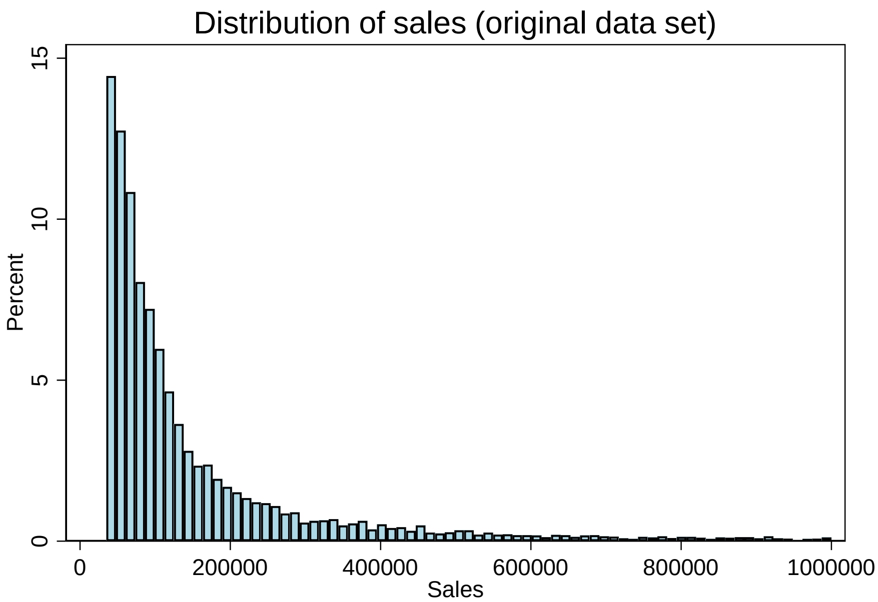
6.3 Sample compared to original data set

In order to draw any conclusions from the result, the sample needs to be representative for Norwegian firms. The industry distribution is representative as we considered this aspect when matching firms. Additionally, we need to evaluate whether the sample is representative with regards to the firm size. In order to do this, we compare the distribution of two important aspects of the firm; sales, and total assets.

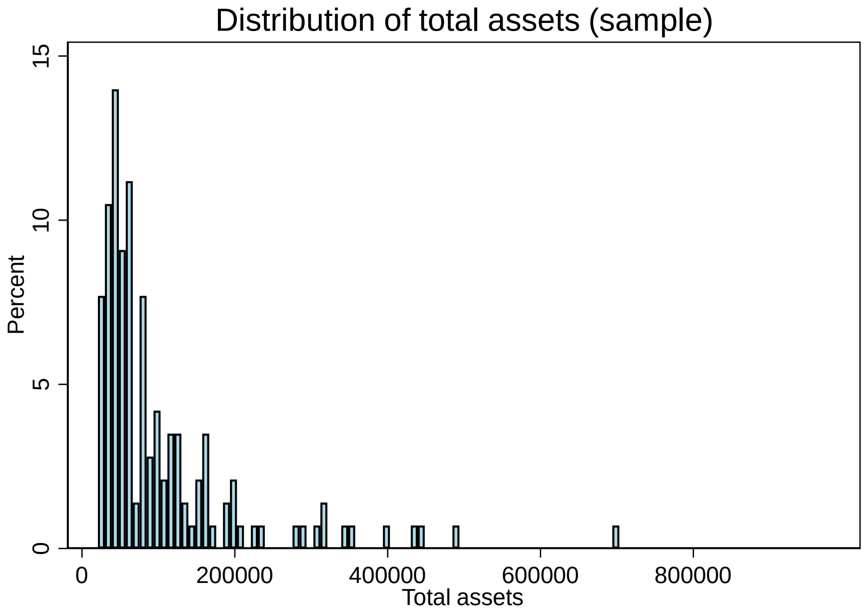
The maximum and minimum limits for sales revenue is the same for both our sample and the database. The minimum limit is as mentioned NOK 35 million, while the maximum limit is NOK 1 000 million. The maximum limit removes outliers, which is beneficial as they distort the distribution, making the samples difficult to compare. The maximum limit removed approximately 5% of the data.

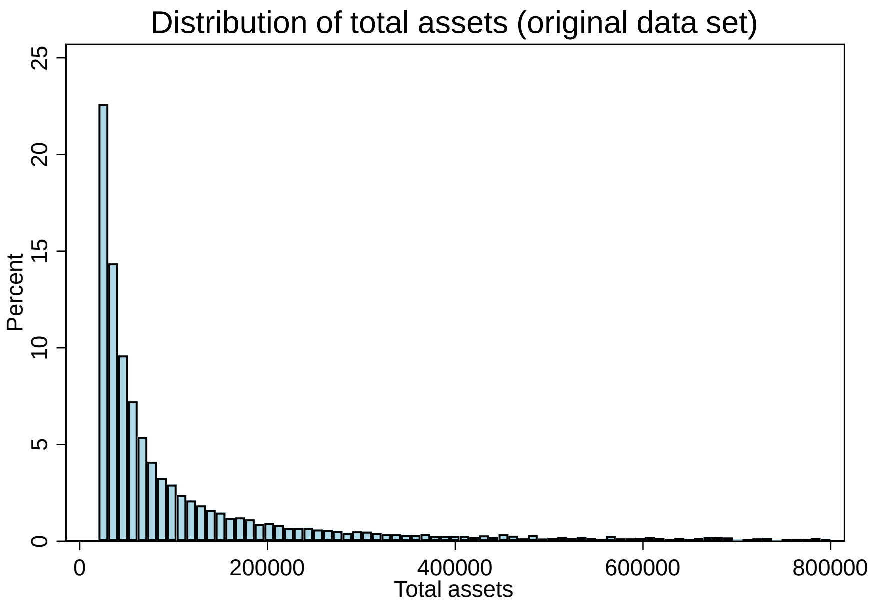
From the graph below, we see that the distribution of firms with regards to sales is relatively similar. Both distributions are highly concentrated around the minimum limit and the proportion of firms decreases as sales increase. The proportion of firms around the peak is similar, both in our sample and the original database.





A similar comparison is made using total assets. The lower limit is NOK 20 million and the maximum limit is NOK 1000 million. The maximum limit removed 8% of the firms. The graph below shows that the distribution for our sample and the original database is relatively similar. Both are concentrated around the lower limit and the proportion of firms decreases as total assets increase. The distribution is similar, but the proportion of firms around the peak is different.





Based on these results we find that the firms included in the sample have a relatively similar distribution to those of the original sample. The industry weight is similar, and the size distribution, with regards to both sales and assets, is relatively similar. We therefore conclude that the sample is a relatively accurate representation of Norwegian firms. However, this is only for Norwegian firms of a significant size. Therefore, the results from this study may not be applicable to very small firms.

6.4 Hold-out-sample

The hold-out-sample was sampled in a similar way to the original sample. However, the restrictions imposed on the sample were less strict. The final ratio profile did not include cash flow figures, which contributed to the hold-out-sample being larger. This resulted in a hold-out-sample of 379 firms, with 219 bankrupt firms and 160 non-bankrupt firms.

The minimum sales and total assets requirements were reduced to NOK 20 million and NOK 10 million respectively. Consistency between the industry distribution in the sample and in the population was not directly considered. However, given that the hold-out-sample is larger than the original sample, the distribution should be close to the ideal distribution as the firms were randomly selected.

One difference between the hold-out-sample and the original sample is that the hold-out-sample is unbalanced. There are more bankrupt firms than non-bankrupt. This is because the issue of outliers was bigger for non-bankrupt firms. We did not examine whether this was an issue related to the random sample of firms or if this was a general issue with the non-bankrupt firms.

6.5 Quality of data

We found the data quality of the database to be inconsistent. One important problem was missing data for important accounting items. Although we were able to gather most of the data needed, some items were incomplete or missing. Some examples are the lack of detailed information with regards to accounts payable/receivable, sales, and liabilities. This problem is particularly evident for bankrupt firms.

Furthermore, there were problems with regards to the point of time of the missing values. For many bankrupt firms, the time span between the last recorded data and the year of bankruptcy was too long. Information close to the date of bankruptcy is crucial, as the models require data one year prior to bankruptcy. In order to mitigate this problem, we used the Forvalt database to manually add the missing information if possible.

After this procedure, a significant number of firms still lacked the financial info one year prior to bankruptcy. This was especially common if the bankruptcy had occurred in the first half of the calendar year. One reason might be that the firms need time to prepare and audit the statements after the turn of the year, with the deadline of submitting being in the summer. If a firm goes bankrupt after the turn of the year but before this deadline, they might be less likely to submit the finical statement. In order to keep the sample at a satisfactory size, we included such firms in the sample of bankrupt firms. However, as the lead time increases, the prediction accuracy is likely to go down. The firms that had a time span over 1.5 years were dropped.

The result of missing information is that the analysis is based on data that is less than optimally adjusted and estimated. This can result in inaccurate models and spurious results. However, considering the relative size of the sample, the effects should be minimal. In some cases, errors caused by suboptimal data may cancel out. We argue that the any limitations and weaknesses associated with the suboptimal accounting data should not take away from the findings to a large extent.

6.6 Limitations of using Norwegian accounting data

A major difference between this study and Altman's original study is that Norwegian accounting data is used, as opposed to American accounting data. This makes the comparison of the results more complicated, as it is problematic to distinguish between the causes of the differences. It is difficult to assess whether the alternative ratios are superior in general, or only for Norwegian companies. This is because Norwegian companies may operate in a business environment different from that in the U.S. The business environment in the U.S. might be significantly different, causing other ratios to be better suited to assess the risk of bankruptcy.

This arguably gives the alternative models an advantage, if the alternative ratios are better suited to assess Norwegian companies. However, this issue is mitigated to some extent as we have re-estimated the coefficient in the Z´´-model using Norwegian data. This enables us to compare the results, but some caution is advised as the scale of the differences might be different.

Comparing our data with Altman's original data does not have any significant merit. This is because the accounting data is based on different countries and from time periods. Therefore, we do not focus on this throughout the paper.

7. Empirical results

In the following section, we discuss the findings of the analyses we have performed. The discussion is based around a select sample of findings. A complete collection of the results obtained from the 86 different models we tested can be found in Appendix 3-11. The results in this section are obtained using IBM SPSS Software.

First, we present and discuss the results from the re-estimated Z´´-model. Next, we present and discuss the effects of substituting one ratio in the Z´´-model with an alternative ratio. This is followed by a presentation and discussion of three alternative models where two ratios are replaced. The alternative models are compared with the Z´´-model with regards to their predictive ability on different samples. Finally, the results are summarized and the best model is suggested.

We also tested models that replaced three or four ratios. The results from these tests were inferior to those models emphasized later in this section (Appendix 5-6). As a result, these models are not discussed in detail.

7.1 The re-estimated Z´´-model

From the studies presented in the literature review, we find that re-estimating the coefficients tends to produce improved results. For this reason, we use the re-estimated Z´´-model as a basis for comparing our alternative models. Using the re-estimated model also ensures that effects from temporal and geographical differences are reduced.

After re-estimating Altman's Z´´- model using our sample, we find that it has an overall accuracy of 77.8%. Type 1 error is 25.3% while Type 2 error is 19%. The model's discriminatory power is significant and the model is able to discriminate between the two groups fairly accurately (Appendix 7). These results are significantly less accurate than the results obtained by Altman in his original study. However, the results are consistent with those obtained by other researchers.

By looking at the F-statistic, we find that every ratio included in the Z´´-model is significant on an individual basis (Appendix 11). This indicates that the ratios contribute to the score of the model, which is also evident from a Wilks' Lambda lower than 1.

From the standardized coefficients, we see that the ratio that most affects the score is EBIT/TA, followed by RE/TA, WC/TA, and BVE/BVL (Appendix 8). EBIT/TA was the most important ratio in the original Z´´-model, similar to our re-estimation. WC/TA was more important than BVE/BVL in the original model, but this is not true for the re-estimation.

Furthermore, RE/TA is more important in the re-estimated model. In this regard, our re-estimated model is partially consistent with Altman's original model.

The re-estimated coefficients are presented in Table 15. An important note from the unstandardized coefficients, is the negative value of WC/TA. According to the model, an increase in WC/TA indicates an increase in default risk. Altman's model uses this ratio as a measure of liquidity where a higher ratio indicates a lower risk of default with regards to liquidity. Hence, there is a lack of consistency between the ratio's reasoning and original results, and the results obtained from this study.

Table 15: Unstandardized Canonical Discriminant Function Coefficients	
WC/TA	-0.966
RE/TA	1.189
EBIT/TA	7.023
BVE/BVL	0.220
(Constant)	-0.094

By looking at the mean value of WC/TA for bankrupt and non-bankrupt firms, we find that it is higher for non-bankrupt firms (Appendix 11). The mean value for non-bankrupt firms is 0.207, while the mean value for bankrupt firms is 0.007. This is consistent with what we would expect. One would therefore assume that the coefficient should be positive, as non-bankrupt firms have a higher average mean value.

An additional re-estimation of the coefficients was done both using a hold-out-sample and a sample of manufacturing firms. This was done solemnly to check whether the issue was consistent. The sample consisting of manufacturing firms is arguably more similar to that used by Altman in his original study, and we should thus expect more similar coefficients. The coefficient was negative for both the hold-out-sample and manufacturing sample.

Other researchers have encountered similar problems. However they do not comment on them to any significant extent. A study by Grice and Ingram (2001) obtained a coefficient of -0.301 when re-estimating Altman's coefficients. They also obtained a coefficient of -0.386 when re-estimating using a sample of manufacturing firms. Moyer (1977) obtained a coefficient of -0.006 using a direct discriminant approach and -0.02 using a step-wise approach. Lack of logical consistency is an issue with the other ratios as well. Boritz et al. (2007) obtained negative coefficients for RE/TA and MVE/BVL.

This exemplifies one of the weaknesses associated with using statistical models to assess the risk of bankruptcy. The method estimates the coefficients that best distinguish between the groups, disregarding logical consistency.

The results obtained from this analysis suggest that Altman's Z´´-score model still offers some informational value with regards to bankruptcy prediction. However, the model seems less suitable today. This is evidenced by the weaker results and the counterintuitive sign of WC/TA.

7.2 The effect of changing one ratio in the Z´´-model

To evaluate whether the alternative ratios are able to improve the model, we replace the corresponding ratio from the Z´´-model with one alternative ratio. The alternative ratios were presented in Table 13. The effects of changing one ratio in the Z´´-model are summarized in Table 16. The coefficients are re-estimated every time a ratio is replaced.

The Z´´-model's results are presented in the first line in the table. They serve as a base with which we can easily compare the effects of replacing one ratio. The effects are discussed below, starting with WC/TA followed by RE/TA, EBIT/TA and BVE/BVL.

Table 16: Effect on accuracy of replacing one ratio in the Z´´-model

Var 1	Var 2	Var 3	Var 4	Overall%	B%	NB%
WC/TA***	RE/TA***	EBIT/TA***	BVE/BVL***	77.8	74.7	81.0
CFO/I	RE/TA***	EBIT/TA***	BVE/BVL***	79.7	77.2	82.3
EBITDA/I*	RE/TA***	EBIT/TA***	BVE/BVL***	79.7	77.2	82.3
EBIT/I***	RE/TA***	EBIT/TA***	BVE/BVL***	81.0	78.5	83.5
WC/TA***	ANI/ATA***	EBIT/TA***	BVE/BVL***	77.8	74.7	81.0
WC/TA***	AEBITDA/AOA***	EBIT/TA***	BVE/BVL***	79.1	77.2	81.0
WC/TA***	AEBIT/AOA***	EBIT/TA***	BVE/BVL***	77.8	75.9	79.7
WC/TA***	RE/TA***	EBITDA/IC*	BVE/BVL***	79.1	81.0	77.2
WC/TA***	RE/TA***	EBIT/IC*	BVE/BVL***	79.1	81.0	77.2
WC/TA***	RE/TA***	EBITDA/OA***	BVE/BVL***	82.3	87.3	77.2
WC/TA***	RE/TA***	EBIT/OA***	BVE/BVL***	80.4	81.0	79.7
WC/TA***	RE/TA***	EBIT/TA***	FA/L***	82.9	83.5	82.3
WC/TA***	RE/TA***	EBIT/TA***	FA/CL***	80.4	78.5	82.3
WC/TA***	RE/TA***	EBIT/TA***	BVEG/BVL***	77.8	74.7	81.0
WC/TA***	RE/TA***	EBIT/TA***	FA/WCF12	75.9	69.6	82.3

(*) significant at 10% sig. level, (**) significant at 5% sig. level, (***) significant at 1% sig. level

Exchanging WC/TA with any of the three alternative ratios results in an improvement in overall accuracy. However, only EBIT/I is significant at the 5% significance level, while EBITDA/I is significant at the 10% significance level. CFO/I is not significant at any level.

The largest change in accuracy is caused by replacing WC/TA with EBIT/I, which results in an overall improvement of 3.2 percentage points. The change is caused by a reduction in Type 1 and Type 2 errors of 3.8 and 1.3 percentage points. The average overall improvement when replacing WC/TA with an alternative ratio is 2.3 percentage points, driven by a reduction in Type 1 and 2 errors of 2.9 and 1.7 percentage points.

Note that CFO/I and EBITDA/I both contain net investments in operational assets. We find that the ratios containing cash flows performed worse than the alternative ratios that contained accounting figures. This is arguably because the fluctuations in cash flows negatively affect the results.

The effect of replacing RE/TA with any of the corresponding alternative ratios is marginal. However, all the ratios are significant at the 5% significance level. The greatest effect is obtained by replacing RE/TA with AEBITDA/AOA. The effect is however relatively small as it only improves overall accuracy by 1.3 percentage points, which is driven by a reduction in Type 1 error of 2.5 percentage points. Using ANI/ATA results in no changes, while using AEBIT/AOA reduces Type 1 errors but increases Type 2 errors.

Replacing EBIT/TA with any of the four alternative ratios improves the overall accuracy. EBITDA/OA and EBIT/OA are significant at the 5% significance level, while EBITDA/IC and EBIT/IC are significant at the 10% significance level. The greatest effect is observed when using EBITDA/OA. The overall improvement is 4.5 percentage points, which is caused by a decrease in Type 1 error of 12.6 percentage points and an increase in Type 2 error of 3.8 percentage points. The average improvement in overall accuracy is 2.4 percentage points. The average reduction in Type 1 error is 7.9 percentage points, while the average increase in Type 2 error is 3.2 percentage points.

Substituting BVE/BVL with one of the four alternatives produces mixed results. Using FA/L and FA/CL improves the model, while using BVEG/BVL has no effect, and FA/WCF12 decreases the overall accuracy. All coefficients are significant at the 1% significance level, except FA/WCF12 which is not significant. FA/L has the greatest effect on the model of all ratios. It improves the overall accuracy with 5.1 percentage points, as a result of a reduction in Type 1 and 2 errors of 8.8 and 1.3 percentage points respectively. The average overall improvement is 1.5 percentage points. The average reduction in Type 1 and 2 errors is 1.9 and 1.0 percentage points.

On average, replacing one ratio resulted in an increase in overall accuracy of 1.7 percentage points. The main driver is the average decrease in Type 1 errors of 3.7 percentage points, while Type 2 errors increased by 0.4 percentage points. The majority of ratios were significant when used as replacements. However, only three ratios were significant when a Chi-squared difference test was performed (Appendix 3). This indicates that the majority of the ratios only offer an improvement of the information contained in the model, but no new information.

Based on these results, we argue that our alternative ratios are able to improve the predictive ability of Altman's model. This supports our hypothesis that ratios picked through deductive reasoning are better suited to predicting bankruptcies than the conventional ratios used by Altman. We argue this is driven by the fact that the selected ratios more accurately measure the aspect they are supposed to evaluate. The reason for this is discussed at length in the Variable Selection section.

7.3 Three alternative models

In the following sub-sections, we present three alternative models to Altman's Z´´-model. The models are based on the Z´´-model, where two ratios are replaced with alternative ratios. In order to develop these models, we tested 86 combinations of Altman's ratios and our alternative ratios.

The different models were selected based on different factors. Model A was selected solemnly on the merit of being the model with the highest overall accuracy in the original sample. Model B was selected as it reduced both Type 1 and 2 errors compared to the Z´´-model. Model C was selected as it had the lowest Wilks' Lambda, hence explaining more of the variation than the other models. Consequently, we include models with a lower overall accuracy if they have other merits. This is something that has not been considered in many bankruptcy studies.

The models are described in greater detail below. These three models are compared with the re-estimated Z´´-score model throughout this section in order to establish which model is the most suited for predicting bankruptcy.

7.3.1 Model A

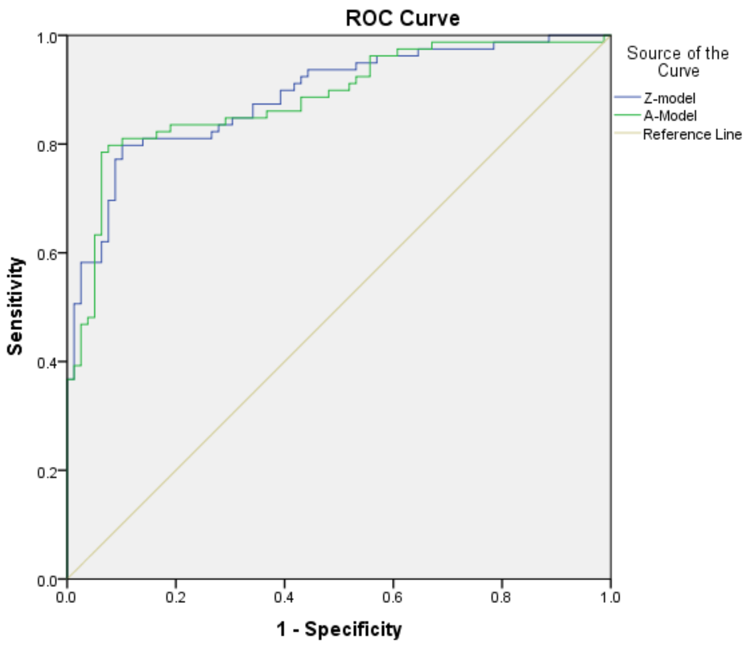
As stated above, Model A is chosen based on having the highest accuracy with regards to the original sample. The model consists of four ratios, two from Altman's Z´´-model and two alternative ratios.

$$Model\ A: \quad -0.60 - 0.72\frac{WC}{TA} + 1.14\frac{RE}{TA} + 2.61\frac{EBITDA}{OA} + 1.10\frac{FA}{L}$$

The overall accuracy of Model A is 86.1%, an improvement of 8.3 percentage points compared to Altman's Z´´-score model (Appendix 4). Type 1 error is down to 6.3%, which is a reduction of 19.0 percentage points compared to the Z´´-model. However, Model A is less accurate when it comes to non-bankrupt firms with a Type 2 error of 21.5%, an increase of 2.5 percentage points. The model's discriminatory power is significant (Appendix 7).

The differences between the two models seem material, and we perform a McNemar test to formally test the significance of the difference. As we have N = 19 discordant pairs, we cannot use normal approximation. As a result, the test is performed using binomial distribution. We find that the test gives a p-value of 0.0044 which is statistically significant at the 1% confidence level (Appendix 10). Thus, we conclude that Model A significantly improves the accuracy compared to model Z´´.

Below we compare the diagnostic ability of Model A to Altman's model, by using a ROC curve. The difference in the AUROC is marginal. Altman's model has an AUROC value of 0.888 while Model A has a value of 0.885. Both of these figures indicate that the models excellently discriminate between the groups (Hosmer & Lemeshow, 2000). The difference between the curves is too small to discriminate between them. It is a strong argument in favor of both models' discriminatory ability to find such a high score.



All ratios are independently significant at the 0.1% confidence level (Appendix 11). The ratio that contributes the most to the score is EBITDA/OA, followed by RE/TA, FA/L, and WC/TA (Appendix 8). We again see that the return on investment measure contributes the most, which is similar to what Altman found in his study. However, WC/TA is the least important ratio in this model, while it was the second most important in the original study. Additionally, RE/TA is the second most important ratio in this model, while it was the least important in Altman's original Z´´-model. This is inconsistent with the results obtained from Altman's original study. However, this is expected as two new ratios are introduced.

Similarly to Altman's re-estimated model, the coefficient of WC/TA is negative. As mentioned previously, this is a counterintuitive interpretation of the liquidity measure. It is important to note that this represents a weakness in the model. This issue is not limited to this model. As mentioned earlier, other researchers have also encountered this problem. This aspect will be evaluated further in the Model Evaluation section.

The improved overall accuracy of the predictions suggest that Model A is superior to the Z´´-model. Although the accuracy with regards to non-bankrupt firms is lower, we argue that the decrease is more than offset by the increase in accuracy for bankrupt firms. Furthermore, the costs associated with Type 1 and Type 2 errors are most likely unequal.

Despite this, there are some arguments in favor of the Z´´-score model. The Z´´-model has a stronger predictive power than Model A when comparing eigenvalues (Appendix 9). The eigenvalue is the ratio between the explained and unexplained variation, where a higher eigenvalue indicates stronger predictive power. Correspondingly, the Wilks' Lambda is smaller for the Z´´-model. This indicates that, although the Z´´-model produces less accurate predictions, it better explains the variance.

The difference in Wilks' Lambda is driven by the removal of EBIT/TA from the model which has the smallest Wilks' Lambda and hence explains the largest degree of the variance. The alternative ratio EBITDA/OA has weaker explanatory power, and the model's total explanatory power, with regards to the variance, is hence decreased. This is offset, to some degree, by the inclusion of FA/L, which has a smaller Wilks' Lambda.

One conclusion that can be drawn is that Model A is more conservative, but not necessarily better at evaluating the risk of default with respect to a company's performance. By conservative we mean that the model classifies firms as bankrupt to a large extent, regardless of the true nature. However, we argue that the superior results outweigh the relatively small difference in explanatory power with respect to variance.

7.3.2 Model B

This model is chosen based on having the highest overall accuracy while reducing both types of errors. Similarly to Model A, the model consists of four ratios, two from

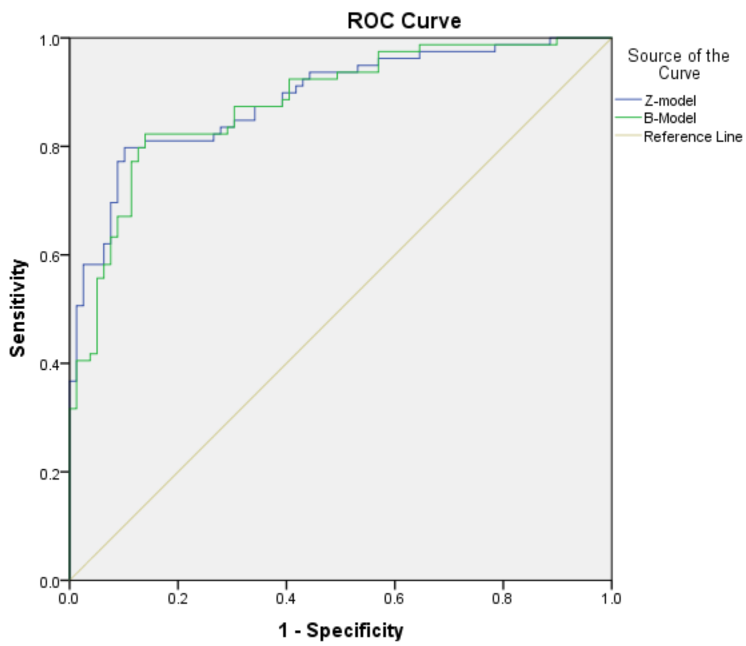
Altman's Z´´-model and two alternative ratios. Note that the FA/L ratio is present in this model as well.

$$Model\ B: \quad -0.38 - 0.38\frac{WC}{TA} + 0.57\frac{ANI}{ATA} + 6.55\frac{EBIT}{TA} + 1.50\frac{FA}{L}$$

Model B has a slightly lower overall accuracy compared to Model A. It is however, more accurate than the Z´´-model, with an overall accuracy of 84.2%. This is 1.9 percentage points lower than Model A, but 6.4 percentage points higher than Altman's model. However, the benefit of this model is that it increases the accuracy for both bankrupt and non-bankrupt firms. Type 1 error for Model B is 13.9% while Type 2 error is 17.7%. This is a reduction of 11.4 and 1.3 percentage points respectively. The predictive power of the model is statistically significant (Appendix 7).

In order to compare the accuracy of the model to the Z´´-model, we perform a McNemar test. Because N = 12 we use a binomial distribution when calculating the p-value of 0.0064 (Appendix 10). This is statistically significant at the 1% confidence level and we conclude that the accuracy is significantly improved by applying Model B.

When comparing the ROC curves of Altman's model and Model B, we find no significant difference between the AUROC. The Z´´-model has the aforementioned value of 0.888, while Model B has value of 0.881. Again, both models are considered to have an excellent discrimination. The values are too similar to accurately discriminate between the models.



The ratios are all independently significant and the ratio with the largest relative contribution is EBIT/TA (Appendix 8 and 11). This is followed by FA/L, WC/TA, and ANI/ATA. The profitability measure is again the most important, just as in Altman's original Z´´-model. ANI/ATA is the least important ratio, just as its corresponding ratio in the original Z´´-model. From this we see that the relative contribution of the ratios in Model B is relatively consistent with the original Z´´-model. As for the previous models, the coefficient of WC/TA is negative.

The improved accuracy of this model suggests it is better suited to estimating distress than Altman's Z´´-score. Despite having a lower overall accuracy than Model A, it still produces fairly accurate results. The advantage of this model is that it improves the accuracy with regards to both bankrupt and non-bankrupt firms. Naturally, this is a tradeoff that depends on how one evaluates the costs associated with each type of error. Nevertheless, it represents a viable alternative to Altman's Z´´-score model.

Furthermore, Model B has a lower Wilks' Lambda than the Z´´-model (Appendix 7). Correspondingly, it also has a larger eigenvalue than the Z´´-model (Appendix 9). This suggests that the model has a stronger explanatory power with regards to the variation. The improvement is a result of replacing RE/TA and BVE/BVL with alternative ratios that both have smaller Wilks' Lambdas.

From this, we can conclude that the model is more accurate at predicting financial distress than the Z´´-model. Additionally, despite being less accurate than Model A, we argue that this is a result of Model A being more conservative, not necessarily more accurate.

7.3.3 Model C

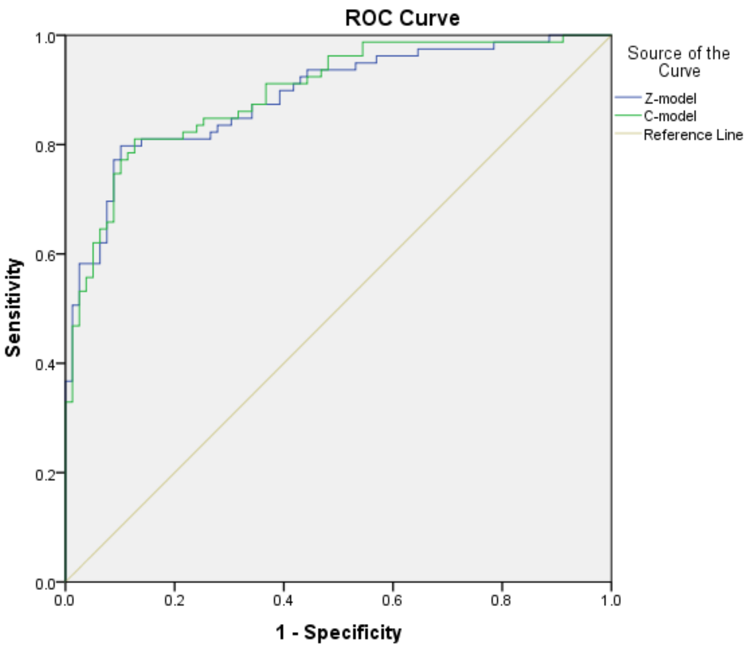
Model C is chosen based on being the model that best describes the variation in the scores, as evidenced by its small Wilks' Lambda and high Eigenvalue. Similarly to Model A and B, the model consists of four ratios, two from Altman's Z´´-model and two alternative ratios. Again we see that FA/L ratio is present in this model.

$$Model\ C\!: \quad -0.36 + 0.67\frac{EBIT}{Interest} + 0.516\frac{RE}{TA} + 5.90\frac{EBIT}{TA} + 0.99\frac{F.Assets}{Liabilities}$$

This model has the lowest overall accuracy of the three alternative models. The overall accuracy is 82.3%, which is an increase of 4.5 percentage points compared to the Z´´-model. This change is driven by a reduction in Type 1 error from 25.3% to 16.5%. Type 2 errors remain unchanged. As for the two other models, Model C is statistically significant (Appendix 7).

Again, we perform a McNemar test using a binomial distribution (Appendix 10). The p-value is 0.0654 and is only significant at the 10% confidence level, when applying traditional confidence levels. This indicates that it is likely that Model C produces materially different results from model Z´´, however this conclusion should be interpreted with a degree of caution.

As for the two previous models, the difference between Altman's model and Model C is too small to provide any material arguments in favor of one model. Model C has the highest AUROC value of all four models, with 0.895. This indicates that the model excellently discriminates between the groups.



All ratios are independently significant, and EBIT/TA is again the ratio that contributes the most to the score (Appendix 8 and 11). This is again followed FA/L. The last two ratios approximately contribute equally to the score. These results are partially inconsistent when compared to those obtained in Altman's original study. One important thing to note is that all the ratio coefficients are now positive. This solves one of the weaknesses of the two former models and our re-estimation of Altman's model; the counterintuitive coefficient of WC/TA.

Considering that the model improves the accuracy for bankrupt firms, while keeping the accuracy for non-bankrupt firms constant, we argue that it is more suited to predict defaults. Although the accuracy is weaker compared to the two other models, it does solve the weakness of negative coefficients. Therefore, it not only represents a viable option to the Z´´-model, it also represents an intuitive option.

Another advantage is that the model has the lowest Wilks' Lambda of all models (Appendix 7) and the largest eigenvalue (Appendix 9). This suggest that Model C is the model with the strongest predictive power as it most adequately explains the variance.

It may be concluded that Model C more accurately evaluates the performance of firms and hence predicts the risk of default with a greater degree of accuracy. Although the accuracy rates from Model C are not as impressive as for Model A and B, they more accurately explain the variance between the groups.

7.4 Failure prediction two years prior

In order to evaluate the alternative models further, we applied the models on the sample with data two years prior to bankruptcy. The discriminant function was not re-estimated. Note that this sample contains 80 bankrupt firms but only 79 non-bankrupt firms. The reason for this is that data two years prior to bankruptcy, but not one year prior, was available for one firm. The results of the predictions are presented below.

Table 17a: Classification Matrix Showing Results Two Years Prior to Bankruptcy

Re-estimated Z´´		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	62	33
	Bankrupt	17	47

Bankrupt hit (%)	58.8
Non-bankrupt hit (%)	78.5
Overall hit (%)	68.6

Table 17b: Classification Matrix Showing Results Two Years Prior to Bankruptcy

Model A		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	53	20
	Bankrupt	26	60

Bankrupt hit (%)	75.0
Non-bankrupt hit (%)	67.1
Overall hit (%)	71.1

Table 17c: Classification Matrix Showing Results Two Years Prior to Bankruptcy			
Model B		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	56	28
	Bankrupt	23	52
Bankrupt hit (%)	65.0		
Non-bankrupt hit (%)	70.9		
Overall hit (%)	67.9		

Table 17d: Classification Matrix Showing Results Two Years Prior to Bankruptcy			
Model C		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	56	28
	Bankrupt	23	52
Bankrupt hit (%)	63.8		
Non-bankrupt hit (%)	73.4		
Overall hit (%)	68.6		

The results indicate that the difference in accuracy between the alternative models and Altman's model is smaller when using accounting figures two years prior. The overall accuracy ranges from 67.9% to 71.1%.

The re-estimated Z´´-model only predicts 58.8% of the bankrupt firms correct. Compared to this, we see that Model A is still more accurate at predicting bankruptcies. The same is the case for Model B and C. However, none of the alternative models are as accurate as the re-estimated Z´´-score model when predicting the state of non-bankrupt firms.

Nevertheless, Model A is still arguably superior to the Z´´-model. Model B is slightly less accurate overall, but compensates by being more accurate when predicting bankrupt firms. Model C is just as accurate overall, but given the assumed higher cost of Type 1 errors we argue that this model does a better job than Altman's. In general, we see that the improved accuracy of the alternative models is reduced when moving one year back.

7.5 Failure prediction for manufacturers

Altman designed the original Z-score model to assess the risk of default for manufacturers. Although we apply the multi-industry Z´´-model, the ratios included were originally selected based on manufacturers. Therefore, it can be argued that the ratios used in the Z´´-model are specifically tailored to predicting distress in manufacturing firms.

In order to evaluate whether our alternative ratios are better than the conventional, we assess the predictive power of the four models on a sub-set of our sample, consisting of manufacturers. The sub-set consists of 44 firms, where 21 are bankrupt and 23 are non-bankrupt.

Table 18a: Classification Matrix Showing Results for Manufacturers

Re-estimated Z´´		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	18	3
	Bankrupt	5	18
Bankrupt hit (%)	85.7		
Non-bankrupt hit (%)	78.3		
Overall hit (%)	81.8		

Table 18b: Classification Matrix Showing Results for Manufacturers

Model A		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	19	1
	Bankrupt	4	20
Bankrupt hit (%)	95.2		
Non-bankrupt hit (%)	82.6		
Overall hit (%)	88.6		

Table 18c: Classification Matrix Showing Results for Manufacturers

Model B		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	17	2
	Bankrupt	6	19
Bankrupt hit (%)	90.5		
Non-bankrupt hit (%)	73.9		
Overall hit (%)	81.8		

Table 18d: Classification Matrix Showing Results for Manufacturers

Model C		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	17	2
	Bankrupt	6	19
Bankrupt hit (%)	90.5		
Non-bankrupt hit (%)	73.9		
Overall hit (%)	81.8		

Model A is more accurate for both bankrupt and non-bankrupt firms. Models B and C increased accuracy when predicting bankrupt firms, but decreased accuracy when predicting non-bankrupt. This resulted in an overall accuracy equal to Altman's model. We see that the Z´´-model perform better when used on manufacturing firms, compared to our original sample. This is consistent with our arguments that the ratios were tailored for this industry.

As a final note, we stress that the differences between the models are marginal. Model A accurately predicts three more firms than the other models. Because the sample is relatively small, this has a large effect on the percentage scores. However, a difference of three hits is not sufficient to draw any definite conclusions.

7.6 Out-of-sample test

In order to test the generalizability of the alternative models we performed an out-of-sample test. It is expected that the prediction accuracy is somewhat lower when performing an out-of-sample test. This is evident from the Z-model's mediocre results when tested on other samples in previous studies (Appendix 1). In accordance with Moyer (1977), the secondary sample should be drawn from a later time than the estimation sample. This is because the aim of the model is to predict the future. However, this is not the case as we did not have access to more recent data.

Two samples were produced, one containing bankrupt firms and one containing non-bankrupt firms. We performed similar adjustments to the samples as previously mentioned. However, the restrictions on sales and total assets were slackened. The new sample contained more variation with respect to size. This arguably has a negative effect on the prediction accuracy. Outlier detection and removal was performed. The outliers are most often a result of poor data quality and are therefore not necessarily representative of the firm. The final bankrupt sample contained 219 firms and the final non-bankrupt sample contained 160.

The three alternative models and Altman's re-estimated Z´´-model were used in the test. The coefficients were not re-estimated.

Table 19a: Classification Matrix Showing Results from Hold-Out-Sample

Re-estimated Z´´		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	131	76
	Bankrupt	29	143
Bankrupt hit (%)	65.3		
Non-bankrupt hit (%)	81.9		

Table 19b: Classification Matrix Showing Results from Hold-Out-Sample			
Model A		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	69	29
	Bankrupt	91	190
Bankrupt hit (%)	86.8		
Non-bankrupt hit (%)	43.1		

Table 19c: Classification Matrix Showing Results from Hold-Out-Sample			
Model B		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	100	57
	Bankrupt	60	162
Bankrupt hit (%)	74.0		
Non-bankrupt hit (%)	62.5		

Table 19d: Classification Matrix Showing Results from Hold-Out-Sample			
Model C		**Observed**	
		Non-bankrupt	Bankrupt
Estimated	Non-bankrupt	136	74
	Bankrupt	24	145
Bankrupt hit (%)	66.2		
Non-bankrupt hit (%)	85.0		

We find that the alternative models continue to outperform the Z´´-model when predicting bankrupt firms. Model A correctly predicts 86.8% of the firms in the hold-out-sample. However, it performs poorly when predicting non-bankrupt firms. Model B also performs poorly when predicting non-bankrupt firms, with only 62.5% accuracy. However, it compensates by being more accurate when predicting bankrupt firms. Model C outperforms the Z´´-model for both bankrupt and non-bankrupt firms.

These results support our previous assessment that Model A is conservative, but not necessarily a good model. Furthermore, it supports our conclusion that Model C is superior with regards to accurately describing the variation in the scores. Model C performs better than the other models in the hold-out-sample, which is expected as it has the lowest Wilks' Lambda. This is an argument in favor of the conclusion that Model C is the most generalizable model.

7.7 Altman's Z´´-score with original coefficients

Throughout this study we have used Altman's ratios with re-estimated coefficients. This is because re-estimating the coefficients allows us to better compare Altman's ratios with the alternatives. As mentioned, it mitigates the changes in the business environment that might affect the suitability of the coefficients. Nevertheless, we also test the accuracy of the Z´´-model using the original coefficients.

The analysis is performed on our original sample and on the hold-out-sample.

Table 20a: Classification Matrix Showing Results using Altman's original coefficients

Z´´-model (original sample)			**Observed**	
			Non-bankrupt	Bankrupt
Estimated	Non-bankrupt		67	23
	Bankrupt		12	56
Bankrupt hit (%)		70.9		
Non-bankrupt hit (%)		84.8		

Table 20b: Classification Matrix Showing Results using Altman's original coefficients

Z´´-model (hold-out-sample)			**Observed**	
			Non-bankrupt	Bankrupt
Estimated	Non-bankrupt		51	25
	Bankrupt		109	194
Bankrupt hit (%)		88.6		
Non-bankrupt hit (%)		31.9		

Using the original coefficients and cut-off score, Altman's model is fairly accurate when applied on our original sample. The overall accuracy is 77.8%, which is the same as the overall rate for the re-estimated model.

However, when applied on the hold-out sample the accuracy is significantly different. The model is very accurate at predicting bankrupt firms, with an accuracy of 88.6%. However, the accuracy for non-bankrupt firms is very low at 31.9%. This indicates that the model is consistently producing scores that are too low, resulting in a large number on bankrupt predictions.

With regards to the accuracy with the original sample, the results are somewhat surprising. One would expect that the re-estimated model is superior, because they are derived using the same sample. However, the performance of the original coefficients is close to that of the re-estimated coefficients. This is in contrast to what most other researchers have found, showing a discrepancy between our results and other researcher's results. However, the general argument that re-estimating the model improves the accuracy is still relevant as many researchers have obtained results that support this. This is also evidenced by the weak performance of the model on our hold-out-sample.

7.8 Summarized results

The results of replacing one ratio with a corresponding ratio, indicate that alternative ratio can improve bankruptcy prediction. A majority of the ratios improved the accuracy of the model and most were significant. This represents evidence in favor of our hypothesis that more sophisticated ratios are superior to conventional ratios, with regards to bankruptcy prediction.

However, we also considered combinations where more than one ratio was replaced. This resulted in three alternative models. The results from these models are summarized in Table 21.

Table 21: Summary of results from different samples

	Re-estimated Z´´-model		**Model A**		**Model B**		**Model C**	
Sample	Type 1	Type 2	Type 1	Type 2	Type 1	Type 2	Type 1	Type 2
Original	25.3%	19.0%	6.3%	21.5%	13.9%	17.7%	16.5%	19.0%
2y prior	41.2%	21.5%	25.0%	32.9%	35.0%	29.1%	36.2%	26.6%
Manufacturing	14.3%	21.7%	4.80%	17.4%	8.50%	26.1%	8.50%	26.1%
Out-of-sample	34.7%	18.1%	13.2%	56.9%	26.0%	37.5%	33.8%	15.0%
Wilks' Lambda (original)	0.592		0.689		0.603		0.592	
Eigenvalue (original)	0.689		0.452		0.659		0.689	
AUROC (original)	0.888		0.885		0.881		0.895	

When assessing the models, emphasis is put on the results from the original sample, the two years prior sample, and the results from out-of-sample. The results from the manufacturing sample are less relevant. This is because the Z´´-model is designed to evaluate companies from several industries, not only manufacturing.

The accuracy is emphasized when assessing the models. However, we also consider the Wilks' Lambda, and consequently eigenvalue, of the models. This is because this gives an indication of the generalizability of the models. The AUROC is not considered as the differences in values are insignificant.

The re-estimated Z´´-model and Model A have a high accuracy with regards to one type of error, and a low accuracy with regards to the other. The estimated Z´´-model has the highest Type 1 error rate in all samples, and a relatively low Type 2 error rate. Model A has the highest Type 2 error rate for all samples but the manufacturing. However, it also has the lowest Type 1 error rate for both the original and hold-out sample. These inconsistent results are undesirable, and lowers the confidence in the predictions. Model B and C are more moderate, producing error rates that are neither the lowest nor the highest.

We argue that Model A is superior to the Z´´-model, as the overall accuracy is better. However, the Wilks' Lambda is the highest for Model A. Still, because the Wilks' Lambda for the Z´´-model is the second highest, this is not dominant.

We regard Model B and C as superior to the aforementioned models. They consistently rank as the second or third most accurate models with regards to both types of error. The exception is that Model B has the lowest Type 2 error rate in the original sample, and that Model C has the lowest Type 2 error rate in the hold-out sample. Furthermore, both models have relatively low values for Wilks' Lambda. Model C have the lowest value, followed by Model B.

When ranking Model B and C, we have to consider several factors. Model B is more accurate with the original sample. However, Model C has a lower Wilks' Lambda and produces the best results in the hold-out sample. These two factors arguably represent the model's generalizability, indicating that the model should perform better outside the estimation sample. Because a bankruptcy prediction model is expected to accurately predict in various samples, this quality is highly regarded.

Additionally, we need to consider the signs of the coefficients. Model B has a negative coefficient for WC/TA. This is counterintuitive and reduces the reliability of the model. Model C is the only model that has intuitive coefficients.

Consequently, because of satisfactory results and intuitive coefficients, Model C is recommended as the best model.

8. Model evaluation

In the following section, we discuss whether the data and models satisfy the assumptions imposed by our statistical approach. We also discuss some issues that are related to the approach. It is important to consider whether the assumptions are met, as it affects reliability of the results.

Many researchers have failed to discuss these issues and this has been widely criticized. Altman did not discuss these issues to a great extent for any of his models. In his original study he points out that the final model was based on its superior accuracy which was caused by intercorrelations of variables (Altman E. , 1968). This is a direct rejection of the assumption of no multicollinearity and reduces the reliability of the results in his study. We want to be certain that our results are reliable and we therefore assess whether our study satisfies the necessary requirements.

8.1 Assumptions with regards to MDA

8.1.1 Sample size

Both equal and unequal sample sizes are acceptable. The smallest sample size n should satisfy the requirement that the number of independent variables should not exceed n – 2. In addition, each group should contain minimum two observations. In our case our sample sizes are equal with N = 158 for all models. The requirements are satisfied for all models. The issue of missing data is irrelevant in this case, as no models contain variables with missing data. From this we see that all models satisfy the assumptions and requirements set by sample size.

8.1.2 Outliers

The analysis is highly sensitive to outliers. Including outliers can impact the significance of variables and the model in general. This can result in spurious results which may not produce consistent predictions. We analyzed the samples for all accounting ratios and removed outlier firms. The effect of this is evident from the distributions (Appendix 13) and group distributions (Appendix 15). There are some outliers in Model A and model C, but these are few and should not have a significant impact on the reliability of the models.

8.1.3 Multivariate normality

MDA assumes that the sample data follows a multivariate normal distribution. If this assumption is violated, the data may still have some discriminating power, however the classification may not be optimal. The analysis is robust to violations of this assumption, if the violations is caused by skewness, not outliers. A conservative estimate is that robustness is expected when the smallest group has at least 20 observations, and there are five or less variables (Tabachnick & Fidell, 2013).

The four models we asses use a combination of the following ratios: WC/TA, RE/TA, EBIT/TA, BVE/BVL, EBIT/I, ANI/ATA, EBIT/OA, and FA/L. There are currently no tests that are feasible for testing the normality of all linear combinations of variables. Hence, testing for multivariate normality is not possible. However, we can test for normality of distribution of each variable. One important note is that normality of variables does not ensure multivariate normality.

Having adjusted for outliers, we perform a Shapiro-Wilk test and produce a histogram of frequency of each observation and a normal Q-Q plot for all ratios (Appendix 14-16).

From the results of the Shapiro-Wilk test only WC/TA, EBIT/TA and EBIT/OA are estimated to be normally distributed (Appendix 12). However, the Shapiro-Wilk test is understood to be highly conservative, and additional analysis is necessary in order to evaluate the nature of the distribution. Based on an analysis of the histograms and Q-Q plots, we evaluate WC/TA, RE/TA, EBIT/TA, ANI/ATA, and EBIT/OA to be close to normally distributed.

Although the other ratios do not seem to be normally distributed, they can still be applied in our analysis. All variable samples produce substantial degrees of freedom which should ensure robustness with respect to normality. There is an absence of significant outliers, which is the major source of sensitivity for discriminant analysis. The non-normality seems to be caused by skewedness, which is perfectly illustrated by FA/L.

An additional reason for why the samples may not exhibit normal distributions is that the data is sorted based on requirements of similarity of sales and total assets. This results in the sample simply representing a section of a possibly normally distributed population. We argue that the distribution of EBIT/I suffers from this. EBIT for companies with similar sales and total assets are naturally similar. When the denominator is relatively small compared to the nominator, these types of similarities result in homogenous ratios.

There have been several studies on the normality of financial ratios. The results of these studies are varying. Deakin's (1976) found only one out of eleven variables to be normally distributed. Ezzamel, Mar-Molinero and Beecher (1987) obtained more mixed results, finding some of the ratios studied to be normally distributed. A study by Jacky So (1987) found that removing outliers could improve the normality of financial ratios. However, many ratios remained non-normal the removal of outliers.

Our results are similar to those obtained by other researchers, or marginally better. It is important to note that this is a result of the removal of outliers. The consequences of this are that the data is closer to being normally distributed. However, this is at the expense of loss of possibly valuable information.

Despite the limitations caused by non-normal variable samples, the analysis should be robust following its equal sample size and large degrees of freedom. The groups satisfy the conservative recommendation of 20 observations and five or less variables. The non-normality seems to be caused by skewedness, not outliers.

Therefore, we argue that all models satisfy the assumption to the degree that the data still has discriminatory power. Nevertheless, one should be cautious when interpreting the significance of the model, and keep in mind that the discrimination may not be optimal.

8.1.4 Homogeneity of Variance-Covariance Matrices

MDA assumes that the variance/covariance matrices are homogenous across groups. If this assumption is not met, the model will tend to classify cases into groups with greater dispersion (Tabachnick & Fidell, 2013). We test for homogeneity of variance-covariance matrices by using the Box's M test and comparing the log determinants of each group.

Following the results from the Box's M test, none of the models have groups with variance/covariance matrices that are homogenous (Appendix 16). However, the Box's M Test is arguably a conservative test which is very sensitive to deviations from multivariate normality. It is often recommended to compare the log determinants of the group covariance matrices. This is especially the case when there is a large amount of observations. The covariance matrices are assumed to be relatively homogenous when the log determinants are approximately equal. From our comparison of log determinants, we find that they are relatively equal (Appendix 17). Therefore, we argue that all four models satisfy the criteria of homogeneity.

An additional argument in favor of the models is that large and equal sample sizes usually are robust to violations of this assumption (Tabachnick & Fidell, 2013). Hence, even if the assumption is violated, the analysis is robust and the results can still be interpreted with some degree of confidence.

8.1.5 Absence of multicollinearity

When correlation between two variables is high, the information provided by the variable becomes redundant. This is important as a step in MDA is to invert the variance/covariance matrix. If some of the variables are redundant, the matrix cannot be inverted. If there is a case of collinearity, the solution could be to delete the redundant variables. The covariance and correlation of the variables present in each model are evaluated to check for multicollinearity. We perform this test by producing pooled within-groups matrices and checking if correlation coefficients are larger than 0.8.

No models have ratios with a correlation coefficient of 0.8 or higher (Appendix 18). The highest correlation coefficient is present between WC/TA and RE/TA, at 0.638. Considering none of the matrices show any substantial correlations between the ratios, we argue that multicollinearity is not a problem in any of the models.

8.2 Miscellaneous

8.2.1 Negative WC/TA coefficients

As observed in the models A, B, and Z´´, the coefficient of the ratio WC/TA is negative. This is a counterintuitive interpretation of the ratio, which is supposed to measure liquidity. There are a few reasons as to why the ratio could be negative. These reasons are discussed below.

First, the ratio might be negative due to the method applied when estimating the function that best discriminates between bankrupt and non-bankrupt firms. The method simply assigns coefficients to the given ratios that maximize the difference between the groups. It does not take into consideration that the coefficients need to be logical. In this case, the model is limited due to the statistical properties of discriminant analysis. Therefore, we argue that there is nothing inherently problematic with using WC/TA. We consider this to be the most likely reason for the negative coefficient.

On the other hand, the counterintuitive sign might be driven by aspects other than the statistical properties of discriminant analysis. If this is the case, the use of WC/TA in bankruptcy prediction should be avoided. This could for instance be caused by the ratio having a relatively large variation. The issue of negative coefficients is persistent in three different models that we produced. It was also present in two different samples and one sub-sample that we applied. Additionally, it has been present in different studies using samples from different countries. Taking into consideration the relative prevalence of the issue, we argue that it is somewhat likely that the negative coefficient is caused by the inherent nature of the ratio. If this is the case, researchers should reconsider using WC/TA in bankruptcy prediction models, based on discriminant analysis.

Considering that the mean value for WC/TA for non-bankrupt firms is larger than that of bankrupt firms, it is hard to argue that the negative coefficient is intuitive in any sense. However, if this was not the case in our sample, the negative sign could have been appropriate. There are some reasons for why non-bankrupt firms might have smaller WC/TA ratios.

First, there is a possibility that successful companies have a relatively low ratio of current assets to total assets, compared to distressed companies. This would explain why having a high WC/TA ratio is considered negative, as working capital is defined as current assets minus current liabilities. However, this is unlikely, as the main driver behind the ratio of current assets to total assets is arguably industry.

Second, it is possible that successful firms finance more of their current assets through debt. This might be the case as stable and successful companies are able to take up new debt with better terms. Meanwhile, bankrupt firms might not have the possibility to take up favorable debt and are hence forced to finance current assets through equity. This is more reasonable, however it is unlikely that firms that default do not have relatively large current liabilities.

Additionally, larger firms are arguably less likely to default than smaller firms. If working capital makes up a relatively small value for all firms, the denominator favors smaller firms. As stated in the criticism of Altman's ratios, using total assets as the denominator implicitly states that smaller firms are less likely to default. This is somewhat reasonable as it is a result of the choice of denominator, not any inherent aspects of the firms.

The illogical coefficient is further evidence of why this ratio is unsuited to represent the liquidity situation of a firm. However, the measure is present in both the best and second-best model, with regards to accuracy in the original sample. This decreases the reliability of the models and the results they achieve. A comparison can still be made between these models and the Z´´-model as they all contain the ratio, and the spurious results caused by it are thus held constant.

8.2.2 Limitations of MDA

The goal of this study is not to produce the best bankruptcy prediction model, but rather to test whether more sophisticated accounting ratios are better at predicting bankruptcies. Because Altman's Z´´-model was chosen as a benchmark, it was necessary to perform an analysis using the same methods. There are several weaknesses with using discriminant analysis compared to alternative methods. These weaknesses and alternative methods are briefly discussed below.

First, the assumption of multivariate normality is often violated. As mentioned, Deakin (1976) found that out of eleven popular ratios, only one was arguably normally distributed. This could lead to biased significance tests. The assumption of equal covariance-variance matrices is often violated as well resulting in biased significance tests (Balcaen, 2006). The tendency to violate two key assumptions reduces the reliability of discriminant analysis as a tool for predicting bankruptcies. Although large samples with few variables often are robust to these violations, it still represents a major weakness.

Furthermore, it is difficult to estimate prior probabilities of bankruptcy and the cost of each type of error. These values should however be estimated as they are important for estimating the optimal cut-off score. The difficulty of estimating these values is a limitation, but it can be partially accounted for by using a range of cut-off scores and performing scenario analyses.

There are several alternatives to MDA, such as logit models, probit models, hazard models, and neural networks. Given the tendency of discriminant analysis to violate its assumptions, the generalizability of the models might be unreliable. Hence, the relevant alternative methods might be better options in some situations. Again, we stress that these methods are outside the scope of this paper.

9. Conclusions and suggestions for future research

Bankruptcy prediction has been a widely discussed topic in the last century. Through the years, a variety of models have been developed. However, our focus has been put on the predictive ability of the ratios included. Throughout this study, our aim has been to provide an answer to our hypotheses; whether modern ratios based on deductive reasoning are better suited to predict bankruptcy than conventional ratios. By using Altman's Z´´-score model as a benchmark, we have also tested the generalizability of the Z´´-score model.

We studied a sample of Norwegian firms over the period 2007 to 2015. By doing so, we had both temporal and geographical differences compared to Altman's model. Some of these differences are mitigated by re-estimating the coefficients of the model. However, these differences might make the comparison of the results inaccurate.

Our initial analysis, replacing only one ratio at a time, provided significant evidence in favor of our hypothesis. Ten out of fifteen alternative ratios improved the overall accuracy of Altman's model when replacing the corresponding ratio. Four ratios produced identical results, while only one reduced the accuracy. This supports our first hypothesis that the alternative ratios are superior to the conventional ratios.

Having tested 86 different combinations of ratios, we chose three models to be tested thoroughly. These models include two of the original ratios and two of our alternative ratios. Although only three models were chosen, a majority of models using combinations of two or more alternative ratios improved the overall accuracy of Altman's model. Furthermore, many of the models were also better at explaining the variation. This supports our second hypothesis that the alternative ratios are able to produce better bankruptcy prediction models.

The prediction accuracy of the three models, applied on the samples of bankrupt firms, are consistently better compared to Altman's model. With regards to the prediction of non-bankrupt firms, the improvements are not significant. Although the alternative models improve the accuracy when using our samples, the accuracy is worse compared to that obtained by Altman using his original sample. However, this is consistent with the findings of several other studies that re-estimated Altman's coefficients (Appendix 2).

When testing the models with hold-out samples, the results are inconsistent. All models predict bankrupt firms more accurate than the re-estimated Z´´-model. However, only Model C predict non-bankrupt firms more accurate than the benchmark. Ideally, we would test the models using different hold-out samples from more recent times. This is because the aim of the model is to make accurate predictions. However, more recent data was not available.

With regards to the underlying assumptions, MDA has been criticized for being strict. However, researchers have found that the technique can tolerate some deviations, and argue that the accuracy is a good indicator of satisfaction (Klecka, 1980). Our tests show that the models satisfy the assumptions to some degree. The accuracy is good, and because the model has a large, balanced sample it is robust to minor violations. Thereby, we are fairly confident with the conclusions drawn from the results, although some discretion is necessary. It is important to note that his problem also is present in many other bankruptcy prediction studies (Balcaen, 2006).

From our findings, we conclude that Model C is the superior model. It produces consistent results and is the most accurate using the hold-out sample, and it also has the lowest Wilks' Lambda. Additionally, it replaces the most criticized ratio in our evaluation, namely WC/TA. It also includes the most used alternative ratio, FA/L, which was present in all three alternative models and had the largest effect on an individual basis. This supports our second hypothesis as we are able to produce model using alternative ratios that is superior to the Z´´-model.

We consider the results obtained in this study to be of value to a firm's stakeholders, particularly investors and lenders. Our findings emphasize the importance of thoroughly assessing the ratios that are applied. A ratio that we do not recommend using in the future, based on our results, is WC/TA. On the other hand, we highly recommend using FA/L when assessing the risk of default. As a final note, we want to stress that today's business environment is more heterogeneous than ever. This means that it is difficult to produce a model that accurately predicts the risk of default for all firms, as evidenced by the failing accuracy of older models. Therefore, we recommend using or producing specialized models when necessary.

9.1 Suggestion for future research

We highly recommend other researchers to perform a similar analysis on samples of firms from outside of Norway. Additionally, we recommend performing a similar analysis using other well-known models, such as Ohlson's model. If possible, researchers should consider normalizing the sample data. Alternative variables that should be specifically considered are Interest Coverage ratios and Financial Assets/Liabilities as they were important variables in this study.

The alternative variables presented in this study, and other similarly sophisticated variables, may also be applicable for research not regarding bankruptcy prediction. We therefore advice other researchers to consider applying some less conventional ratios when performing their research, especially with regards to performance assessment, stock market predictions and comparisons between companies.

References

Almamy, J., Aston, J., & Ngwa, L. N. (2016, February). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance, 36*, 278-285. Retrieved from http://www.sciencedirect.com/science/article/pii/S0929119915001583

Altman, E. (1968, September). Financial Ratios, discriminant analysis and the prediction of corporate bankruptcy.

Altman, E. (1983). *Corporate Financial Distress: A complete guide to predicting, avoiding and dealing with bankruptcy.* John Wiley & Sons.

Altman, E. (1984, June). The success og business failure prediction models: An international survey. *Journal of Banking and Finance*(8), pp. 171-198.

Altman, E. (2000, July). *Predicting Financial Distress of Companies: Revisiting the Z-score and Zeta Models.* New York University, Stern School of Business.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance, 1*(1), 29-54. Retrieved 8 21, 2017, from https://econpapers.repec.org/article/eeejbfina/v_3a1_3ay_3a1977_3ai_3a1_3ap_3a29-54.htm

Altman, E., & Lavallee, M. (1981). Business Failure Classification in Canada. *Journal of Business Administration*.

Altman, E., Baidya, T., & Riberio-Dias, L. (1979). Assessing Potential Financial Problems of Firms in Brazil. *Journal of International Business Studies*.

Avenhuis, J. O. (2013, 11). Testing the generalizability of the bankruptcy prediction models of Altman, Ohlson and Zmijewski for Dutch listed and large non-listed firms. Twente, Netherland. Retrieved 10 30, 2017, from http://essay.utwente.nl/64326/1/MSc_Oude%20Avenhuis.pdf

Balcaen, S. (2006, March). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting Review, 38*(1), 63-93. Retrieved 11 9, 2017, from http://www.sciencedirect.com/science/article/pii/S0890838905000636#bib69

Bandyopadhyay, A. (2006). Predicting probability of default of Indian corporate bonds: logistic and Z-score model approaches. *The Journal of Risk Finance, 7*(3), 255-272.

Retrieved 10 30, 2017, from http://www.emeraldinsight.com/doi/full/10.1108/15265940610664942

Beaver, W. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*(4), pp. 71-111.

Begley, J., Ming, J., & Watts, S. (1996, December). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies, 1*(4), 267-284. Retrieved 10 30, 2017, from https://link.springer.com/article/10.1007%2FBF00570833?LI=true

Berner, E., & Olving, M. (2013). *Utbredelse og virkninger av IFRS i Norge; en empirisk studie om utbredelsen av IFRS i Norge, og konsekvensen standarden har på rapporterte nøkkeltall.* Bergen. Retrieved 10 13, 2017

Berner, E., Mjøs, A., & Olving, M. (2016). *Norwegian Corporate Accounts - Documentation and quality assurance of SNF's and NHH's database of accounting and company information for Norwegian companies.* Centre for Applied Research at NHH. Retrieved 9 28, 2017

Bernhardsen, E., & Larsen, K. (2007, October). Modelling credit risk in the enterprise sector - further development of the SEBRA model. *Economic Bulletin*, pp. 102-108.

Boritz, E. J., Kennedy, D. B., & Sun, J. Y. (2007). Predicting Business Failures in Canada. *Accounting Perspectives, 6*(2), 141-165. Retrieved 10 20, 2017, from http://onlinelibrary.wiley.com/doi/10.1506/G8T2-K05V-1850-52U4/full

Bruno, B., Tanja, S.-P., & Keglevic, S. (2014). ASSESSMENT OF BANKRUPTCY PREDICTION MODELS' APPLICABILITY IN CROATIA. *An Enterprise Odyssey. International Conference Proceedings*, (pp. 543-561). Zagreb. Retrieved 10 30, 2017, from https://search.proquest.com/docview/1566187204/fulltextPDF/793CF199B3A942E2PQ/1?accountid=37265

Bureau of Economic Analysis. (2017). *Interactive data - National Data*. Retrieved 11 22, 2017, from Bureau of Economic Analysis: https://www.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=3&isuri=1&1910=x&0=-99&1921=survey&1903=1&1904=1936&1905=2017&1906=q&1911=0

Caouette, J. B. (2008). *Managing Credit Risk* (2. edition ed.). Wiley.

Castagna, A. D., & Matolcsy, Z. P. (1981). The Prediction of Corporate Failure: Testing the Australian Experience. *Australian Journal of Management*.

Celli, M. (2015). Can Z-Score Model Predict Listed Companies' Failures in Italy? An Empirical Test. *International Journal of Business and Management, 10*(3), 57-66. Retrieved 10 30, 2017, from http://www.ccsenet.org/journal/index.php/ijbm/article/view/43926/24864

Charitou, A., Neophytou, E., & Charalambous, C. (2004, September). Predicting corporate failure: empirical evidence for the UK. *European Accounting Review*(13), pp. 465-497.

Cheng, A. C., Cheung, J. K., & Gopalakrishnan, V. (1993). On the Usefulness of Operating Income, Net Income and Comprehensive Income in Explaining Security Returns. *Accounting and Business Research, 23*(91), 195-203.

Christopoulos, A., Gerantonis, N., & Vergos, K. (2009). Can Altman Z-score model predict business failures in Greece? In C. C. Frangos (Ed.), *Quantitative and qualitative methodologies in the economic and administrative sciences,* (pp. 149-154). Athens. Retrieved 10 25, 2017, from https://books.google.no/books?hl=no&lr=&id=ri9ZK2nvl64C&oi=fnd&pg=PA149&dq=Altman+z+score+&ots=g-hVcpAaoc&sig=ojnMgde4RtvIBRk4HwuvEQ4cUZg&redir_esc=y#v=onepage&q=Altman%20z%20score&f=false

Deakin, E. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research, 10*, 167-179. Retrieved 8 22, 2017

Deakin, E. B. (1976, January). Distributions of Financial Accounting Ratios: Some Empirical Evidence. *The Accounting Review, 51*(1), 90-96. Retrieved 9 11, 2017, from http://www.jstor.org/stable/pdf/245375.pdf

Deloitte. (n.d.). *Use of IFRSs by jurisdiction - G20 domestic listed companies*. Retrieved 11 22, 2017, from UK Accounting Plus: https://www.iasplus.com/en-gb/resources/ifrs-topics/use-of-ifrs-g20#totals

Eurostat. (2017). *Real GDP growth rate - volume*. Retrieved 11 21, 2017, from Eurostat: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tec00115&plugin=1

Ezzamel, M., Mar-Molinero, C., & Beecher, A. (1987). On The Distributional Properties of Financial Ratios. *Journal of Business Finance & Accounting, 14*(4), 463-481. Retrieved 11 9, 2017, from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=1&sid=1d097b8e-e5e6-4be4-9fbf-0de48501b470%40sessionmgr4006

Fagerland, M., Lydersen, S., & Laake, P. (2013, July). The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology*.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*(27), pp. 861-874.

Field, A. (2009). *Discovering Statistics Using SPSS* (3 ed.). SAGE Publications. Retrieved 11 26, 2017, from http://www.soc.univ.kiev.ua/sites/default/files/library/elopen/andy-field-discovering-statistics-using-spss-third-edition-20091.pdf

Forvalt. (n.d.). *Om Proff Forvalt*. Retrieved 11 10, 2017, from https://www.forvalt.no/Om/OmProffForvalt/om-forvalt.aspx

Gentry, J. A., Newbold, P., & Whitford, D. T. (1985). Classifying Bankrupt Firms with Funds Flow Components. *Journal of Accounting Research, 23*(1), 146-160. Retrieved 10 26, 2017, from http://www.jstor.org/stable/2490911?seq=12#page_scan_tab_contents

Gerety, M., & Lehn, K. (1997). The Causes and Consequences of Accounting Fraud. *Managerial and Decision Economics, 18*(7/8), 587-599.

Gissel, J. L., Giacomino, D., & Akers, M. D. (2007). A Review of Bankruptcy Prediction Studies: 1930-Present. *Journal of Financial Education*, *33*(Winter 2007), 1-42.

Grice, J. S., & Ingram, R. W. (2001, October). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research, 54*(1), 53-61. Retrieved from http://www.sciencedirect.com/science/article/pii/S0148296300001260

Gutzeit, G., & Yozzo, J. (2011, March). Z-score Performance Amid Great Recession. *American Bankruptcy Institute Journal*(30), pp. 44-45.

Hillegeist, S., Keating, E., Cram, D., & Lundstedt, K. (2004, March). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*(9), pp. 5-34.

Horrigan, J. O. (1968). A Short History of Financial Ratio Analysis. *The Accounting Review, 43*(2), 284-294. Retrieved 10 8, 2017, from http://www.jstor.org/stable/pdf/243765.pdf

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2 ed.). Wiley. Retrieved 11 8, 2017

Jackson, R. H., & Wood, A. (2013, 9). The performance of insolvency prediction and credit risk models in the UK: A comparative study. *The British Accounting Review, 45*(3), 183-202. Retrieved 10 30, 2017, from http://www.sciencedirect.com/science/article/pii/S0890838913000498

Klecka, W. (1980). *Discriminant Analysis.* Sage Publications.

Ko, C. J. (1982). *A Delineation of Corporate Appraisal Models and Classification of Bankrupt Firms in Japan.* New York University.

Lifschutz, S. (2010). Predicting Bankruptcy: Evidence from Israel. *International Journal of Business and Management, 5*(4), 133-141. Retrieved 10 30, 2017, from http://www.ccsenet.org/journal/index.php/ijbm/article/view/4933/4618

Machek, O. (2014). Long-term Predictive Ability of Bankruptcy Models in the Czech Republic: Evidence from 2007-2012. *Central European Business Review*(2), 14-17. Retrieved 10 30, 2017, from https://www.ceeol.com/search/article-detail?id=198216

Manly, B. F. (2004). *Multivariate Statistical Methods: A primer* (3 ed.). CRC Press. Retrieved 11 26, 2017

McNemar, Q. (1947). Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages. *Psychometrika*(12).

Moyer, C. (1977). Forecasting Financial Failure: A Re-examination. *Financial Management, 6*(1), 11-17. Retrieved 10 30, 2017, from http://www.jstor.org/stable/3665489?sid=primo&origin=crossref&seq=1#page_scan_tab_contents

Næss, A. B. (2015, January). Nearest Neighbor Frame Classification for Articulatory Speech Recognition. *Thesis for the degree of Philosophiae Doctor*.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*(18), pp. 109-131.

O'Shaughnessy, P. (2015, 5 19). *The Original Value Factor: Price-to-book*. Retrieved 11 20, 2017, from THe Investor's Field Guide: http://investorfieldguide.com/the-original-value-factor-price-to-book/

Pongsatat, S., Ramage, J., & Lawrence, H. (2004). Bankruptcy Prediction for Large and Small Firms in. *Journal of Accounting and Croporate Governance , 1*(2), 1-13. Retrieved 10 30, 2017, from http://jacg.rd.fcu.edu.tw/dl/1201.pdf

Schermelleh-Engel, K., & Werner, C. (2010). Deciding Between Competing Models: Chi-Square Difference Tests. Retrieved from https://www.psychologie.uzh.ch/dam/jcr:ffffffff-b371-2797-0000-00000fda8f29/chisquare_diff_en.pdf

Sherbo, A., & Smith, A. (2013, December). The Altman Z-score Bankruptcy Model at Age 45: Standing the Test of Time? *American Bankruptcy Institute Journal*(32), pp. 40-41.

Shumway, T. (2001, January). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*(74), pp. 101-124.

So, J. C. (1987). Some Empirical Evidence on the Outliers and the Non-Normal Distribution of Financial Ratios. *Journal of Business & Accounting, 14*(4), 483-496. Retrieved 11 9, 2017, from http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=1&sid=1d097b8e-e5e6-4be4-9fbf-0de48501b470%40sessionmgr4006

Stevens, J. P. (2009). *Applied Multivariate Statistics For The Social Sciences* (5 ed.). Routledge.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6 ed.). Pearson. Retrieved 11 7, 2017

Taffler, R. J. (1983). The Assessment of Company Solvency and Performance Using a Statistical Model. *Accounting and Business Research, 13*(52), 295-308. Retrieved 8 22, 2017, from http://www.tandfonline.com/doi/abs/10.1080/00014788.1983.9729767?journalCode=rabr20

Taffler, R., & Tisshaw, H. (1977). Going, Going, Gone - Four Factors Which Predict. *Accountancy*.

Takahashi, K., Kurokawa, K., & Watese, K. (1979). *Predicting Corporate Bankruptcy through Financial Statements.* Society of Management of Keio University.

The World Bank. (2017). *GDP growth (annual %)*. Retrieved 11 22, 2017, from The World Bank: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=BG-NO

US Bureau of Labour Statistics. (n.d.). US Employment by sector. Retrieved 10 8, 2017

Van Frederikslust, R. (1978). Predictability of Corporate Failure.

Wang, Y., & Campbell, M. (2010). Do Bankruptcy Models Really Have Predictive Ability? Evidence using China Publicly Listed Companies. *International Management Review, 6*(2), 77-82. Retrieved from https://search.proquest.com/docview/751931209/fulltextPDF/5917B186002F4B7CPQ/1?accountid=37265

Weinrich, G. (1978). Prediction of Credit Worthiness, Direction of Credit Operations by Risk Classes.

Yardeni, D., Abbott, J., & Quintana, M. (2017). *Stock Market Briefing: S&P 500 Financial Ratios.* Yardeni Research Inc. Retrieved 11 22, 2017, from http://www.yardeni.com/pub/spxratios.pdf

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*(22), pp. 59-82.

Appendix

Appendix 1 – Collection of previous studies using Z-models

Study	Country	Years	Overall	Bankrupt	Non-Bankrupt
Gutzeit & Yozzo (2011)	US	2004-2009	-	90.0%	70.2%
Grice & Ingram (2001)	US	1988-1991	56.1%	68.2%	54.9%
Grice & Ingram (2001)*	US	1988-1991	87.6%	48.6%	94.9%
Grice & Ingram (2001)m	US	1988-1991	69.1%	69.2%	69.1%
Grice & Ingram (2001)*m	US	1988-1991	86.4%	55.4%	92.1%
Begley, Ming & Watts (1996)	US	1980-1991	78.2%	81.5%	74.9%
Begley, Ming & Watts (1996)*	US	1980-1991	78.4%	78.5%	78.4%
Charles Moyer (1977)	US	1946-1965	75.0%	60.8%	88%
Charles Moyer (1977)*	US	1946-1965	88.1%	95%	82%
Charles Moyer (1977)*	US	1946-1965	90.5%	95%	86%
Boritz, Kennedy & Sun (2007)	Canada	1987-2002	71.8%	78.9%	64.8%
Almamy, Aston & Ngwa (2016)*	UK	2000-2013	54.4%	60.6%	54.0%
Jackson & Wood (2013)*	UK	2000-2009	40.1%	52.0%	39.9%
Jeroen Avenhuis (2013)d*	Netherland	2008-2012	80.6%	35.7%	82.5%
Bruno, Keglevic, Tanja (2014)c	Croatia	2008-2011	80.0%	70.0%	90.0%
O. Machek (2014)	Czech Rep.	2007-2012	44.3%	-	-
Massimiliano Celli (2015)	Italy	1995-2013	87.3%	84.3%	90.1%
Christopoulos, Gerantonis & Vergos (2009)	Greece	2003-2007	56.6%	65.9%	54.2%
Wang & Campbell (2010)	China	1998-2008	51.2%	96.3%	51.1%
Wang & Campbell (2010)*	China	1998-2008	84.7%	85.2%	84.7%
Bandyopadhyay (2006)*	India	1998-2003	83%	82%	84%

Pongsatat, Ramage & Lawrence (2004)a	Thailand	1998-2003	58.9%	90.5%	40.0%
Pongsatat, Ramage & Lawrence (2004)b	Thailand	1998-2003	64.1%	94.9%	16.0%
Shilo Lifschutz (2010)	Israel	2000-2007	62.5%	100%	25%

(*)Re-estimated coefficients, (a)only large asset firms, (b)only small asset firms, (c)small sample size, (d)master thesis, (m) only manufacturing firms

Appendix 2 – List of companies (original sample)

Reg. number	Name	Sector	Year	Bankruptcy
814709442	Novenco AS	Wholesale and retail	2014	0
821465222	Bring Warehousing AS	Transportation and storage	2014	0
829221292	Tema Trading AS	Wholesale and retail	2011	2013
845547432	Solhytten Eiendom AS	Construction	2014	0
853008192	Miras Multimaskin AS	Industry	2010	2012
855543702	Geitanger Bygg AS	Wholesale and retail	2014	0
862957822	Dynatrace AS	Nonclassifiable	2014	0
865863802	Carat Norge AS	Proffesional, scientific and technical services	2014	0
871143072	Veitransport AS	Transportation and storage	2014	0
876810662	Color Print Norge AS	Nonclassifiable	2011	2012
884048842	Logi Trans AS	Transportation and storage	2014	0
886019262	Grove-Knutsen & Co AS	Wholesale and retail	2014	0
887493642	Hellefoss AS	Industry	2011	2013
889099682	RiksTV AS	Nonclassifiable	2010	0
890375472	Tufjordbruket AS	Industry	2014	0
891465432	Longa Industryer AS	Real estate	2011	2012
891704542	Florø Mekaniske Verksted AS	Industry	2014	2015
910115235	Lerøy Alfheim AS	Industry	2014	0
910814958	BRENDE AS	Construction	2014	0
912196593	IDT Automasjon AS	Industry	2015	2017
912904881	SHOE-D-VISION NORGE AS	Commercial services	2014	0
914664292	Canon Norge AS	Wholesale and retail	2014	0
915184979	Peterson AS	Proffesional, scientific and technical services	2010	2012

915428940	Rivenes AS	Transportation and storage	2014	0
916205783	Scana Steel AS	Industry	2013	2015
918352074	Norsea Gas AS	Transportation and storage	2013	0
919145625	Gull-Funn AS	Wholesale and retail	2014	2016
920413382	Pa Consulting Group AS	Proffesional, scientific and technical services	2014	0
921347405	Bema AS	Wholesale and retail	2014	0
923726780	Byggekompaniet AS	Construction	2010	2012
924206527	Takservice AS	Construction	2011	2013
924407573	AS Fiskevegn	Industry	2014	0
928907163	Frekhaug Vinduet AS	Industry	2014	0
929704738	AS Anleggsvirksomhet	Construction	2011	0
929969332	Flisekompaniet AS	Wholesale and retail	2014	0
930694088	Skude Industry AS	Industry	2015	2016
930705462	Oskar og Tormod Wike AS	Construction	2014	0
931510126	Firda Media AS	Nonclassifiable	2014	0
931833340	Black Design AS	Wholesale and retail	2012	2014
933341771	Optical Storage AS	Wholesale and retail	2012	2013
933792110	Hurlum AS	Wholesale and retail	2010	2012
934313038	Carbon Partners AS	Wholesale and retail	2014	0
935429447	Vianor AS	Wholesale and retail	2014	0
935500419	Itab Butikkinnredninger AS	Wholesale and retail	2014	0
935534925	Cavotec Micro-Control AS	Industry	2014	0
935566932	S Sigvartsen Steinindustri AS	Industry	2014	0
935708745	Certex Norge AS	Industry	2014	0
936888739	Blom Geomatics AS	Proffesional, scientific and technical services	2014	0
937076940	Mha Entrepenør AS	Construction	2015	2017
937077467	X-Subsea Norway AS	Nonclassifiable	2013	2015
937567200	Nordlie Auto AS	Wholesale and retail	2014	0
938420718	Ocean Rig AS	Proffesional, scientific and technical services	2014	0
938709599	Norsk Sjømannsforbund	Nonclassifiable	2014	0
938803595	Cecon ASA	Proffesional, scientific and technical services	2013	2015
939483969	Dustin Norway AS	Wholesale and retail	2014	0
940434254	Eiendomsmegler Vest AS	Real estate	2013	0
942269331	JAS Bil AS	Transportation and storage	2012	2014
943387354	Kongsberg Teknologipark AS	Real estate	2014	0
943393222	AS Malmbergs Elektriske	Wholesale and retail	2014	0
943626545	AS MAREX	Industry	2012	0

943659524	Bergen Bunkers AS	Wholesale and retail	2013	2014
944080171	Holtet Pukk & Betong AS	Industry	2014	0
945486260	Adas AS	Industry	2014	2015
946333611	H.J. Økelsrud AS	Wholesale and retail	2014	2016
946506370	New Wave Norway AS	Wholesale and retail	2014	0
947715259	Fosdalen Industrier AS	Industry	2010	0
948640147	CG Glass AS	Industry	2013	2015
948825716	Istrail AS	Industry	2014	0
951118141	Møre Trafo AS	Industry	2014	0
951667595	SN Bygg AS	Construction	2013	2015
952228609	Backe Trondheim AS	Construction	2014	0
952720600	Multibygg AS	Construction	2012	0
954165892	VISMA Retail AS	Nonclassifiable	2014	0
956238900	Modulvegger AS	Industry	2014	0
958029438	Møre Seasfood AS	Wholesale and retail	2012	2013
959299137	Vivo Bokhandel AS	Wholesale and retail	2014	2016
959566704	Enghav AS	Wholesale and retail	2012	2013
960116712	Norske systemarkitekter AS	Nonclassifiable	2011	2012
962276717	Roald & Sønn AS	Wholesale and retail	2014	0
963193149	Modena Fliser AS	Wholesale and retail	2014	0
963979703	Institutt for prosjektledelse AS	Proffesional, scientific and technical services	2011	2013
964725799	Slagen elektro AS	Construction	2014	0
966618841	NEL Hydrogen AS	Industry	2011	2013
968247379	Radøygruppen AS	Industry	2013	0
968469940	Wirtgen Norway AS	Wholesale and retail	2014	0
968586238	Bergersen Flis AS	Wholesale and retail	2014	0
968698907	Totaltek Tekniske AS	Construction	2012	2013
970968857	KIME - Maskinentrepenør Kåre Isaksen AS	Construction	2013	2015
971141603	Lofotprodukt AS	Industry	2013	0
971235845	VRS Installasjon AS	Construction	2012	2014
971579412	Glitter AS	Wholesale and retail	2014	0
974278987	Vero Holding AS	Real estate	2014	2015
974536269	Hellvik hus Flekkefjord AS	Construction	2014	0
974536315	Swets Information Services AS	Wholesale and retail	2013	2014
974788470	Norwater AS	Industry	2014	0
975950247	Walde Gruppen AS	Real estate	2012	2013
975960501	Anlegg Øst AS	Construction	2010	2012
976090993	Taraldset eiendom AS	Construction	2014	0
976114396	AS Vikan Betong	Industry	2013	0
976256697	VITPRO AS	Wholesale and retail	2013	2015
976746996	Nordbohus Romerike AS	Construction	2014	0
979729332	Grunnaleite & Lindstrøm Prosjektservice AS	Construction	2014	0
980488683	Carboline Norge AS	Industry	2014	0

981095332	TF Anlegg AS	Transportation and storage	2014	2015
981126122	Borge rør AS	Construction	2014	0
981682955	Nordfjord kjøtt slakt AS	Industry	2014	0
982031834	Live Nation Norway AS	Proffesional, scientific and technical services	2014	0
982702887	Zacco Norway AS	Commercial services	2014	0
982800781	Andersens Emballasjone & Design AS	Wholesale and retail	2012	2013
983468756	Detaljpartner AS	Wholesale and retail	2011	2013
983522440	Profitek AS	Construction	2011	2013
983668941	NB Marine AS	Industry	2012	2014
983792324	John Galten AS	Construction	2014	0
983805183	Obas øst AS	Construction	2014	0
983979122	Vard Piping AS	Industry	2014	0
983982050	Juvelen Norge AS	Wholesale and retail	2014	2016
984039557	Bergen Group Skarveland AS	Industry	2014	2015
984089198	Rogaland Tekstil AS	Nonclassifiable	2012	2013
984357494	Haukelifjell utvikling AS	Construction	2009	0
984636318	Devold of Norway AS	Industry	2011	0
985090181	Raufoss metall AS	Industry	2010	0
985205027	Vedal Prosjekt AS	Proffesional, scientific and technical services	2014	0
985386854	HAB Construction AS	Construction	2014	0
985445095	Horten Hus AS	Industry	2013	2015
986015248	Callenberg AS	Industry	2014	0
986122230	Nli Odda AS	Industry	2015	2016
986281797	Nor-Reg Systems AS	Industry	2014	2015
986522387	Strukton Rail AS	Construction	2010	2012
986683801	Ck retail AS	Wholesale and retail	2014	2016
986750452	Masai Scandinavia AS	Commercial services	2010	2012
986912827	Stjern Entreprenør AS	Construction	2014	0
987602910	Bilsentergruppen AS	Proffesional, scientific and technical services	2012	2014
987740353	Moods of Norway AS	Wholesale and retail	2016	2017
988193496	Panorama Gruppen AS	Proffesional, scientific and technical services	2014	2016
988225134	AIT Otta AS	Industry	2011	2012
988350877	Marwin Mekaniske AS	Industry	2014	2015
988366889	JK Entrepenør AS	Construction	2011	2012
988423122	Målselv Utvikling AS	Construction	2011	2013
989034138	Rune Øvergård AS	Construction	2010	2012
989396617	Mudenia Elektro AS	Construction	2012	2014
989676989	Arctic Seaworks AS	Construction	2014	2016
989833715	Senterbok Holding AS	Wholesale and retail	2011	2013

990628130	Teknobygg Entreprenør AS	Construction	2011	2012
990637806	Lonnheim Stal AS	Industry	2015	2017
991191852	Destia Norge AS	Construction	2010	2012
991408983	Elvenes Maskin AS	Construction	2015	2017
991786945	Lam Invest AS	Real estate	2011	2013
991893407	Sensonor Technologies AS	Industry	2010	2012
992047453	Miras Vedlikehold og Modifikasjon AS	Industry	2010	2012
992289430	SIVA Shipping Oslo AS	Transportation and storage	2012	2014
992317337	Norsk Kulde Finnsnes AS	Construction	2011	2013
992525150	SIC Processing AS	Nonclassifiable	2011	2012
992684941	Atlantic Offshore AS	Transportation and storage	2014	2016
993405566	Nli Subsea Service AS	Industry	2015	2016
993451207	Norstec AS	Commercial services	2014	2016
994301934	FMV Holding AS	Proffesional, scientific and technical services	2014	2015
998092558	NLI Larvik AS	Industry	2014	2016
998509300	Herøya Industripark AS	Real estate	2014	0

Appendix 3 – Chi-square difference test

Restricted model	**Additional variable**	χ^2_{diff}	df_{diff}	**p-value**
Z´´-model	CFO/I	-0.020	1	0.888
Z´´-model	EBITDA/I	-0.307	1	0.580
Z´´-model	EBIT/I	-3.248	1	0.072*
Z´´-model	ANI/ATA	-1.066	1	0.302
Z´´-model	AEBITDA/AOA	0.179	1	0.672
Z´´-model	AEBIT/AOA	0.210	1	0.647
Z´´-model	EBITDA/IC	0.243	1	0.622
Z´´-model	EBIT/IC	0.245	1	0.621
Z´´-model	EBITDA/OA	-0.336	1	0.562
Z´´-model	EBIT/OA	-0.056	1	0.813
Z´´-model	FA/L	-6.781	1	0.009***
Z´´-model	FA/CL	-2.089	1	0.148
Z´´-model	BVEG/BVL	-1.251	1	0.263
Z´´-model	FA/WCF12	-0.590	1	0.442

(*) significant at 10% sig. level, (**) significant at 5% sig. level, (***) significant at 1% sig. level

Appendix 4 – Prediction accuracy changing two ratios

Model	**Var 1**	**Var 2**	**Var 3**	**Var 4**	**Overall%**	**B%**	**NB%**
Altman	WC/TA	RE/TA	EBIT/TA	BVE/BVL	77.8	74.7	81.0
#15	EBITDA/I	ANI/ATA	EBIT/TA	BVE/BVL	79.1	77.2	81.0
#16	EBITDA/I	AEBIT/AOA	EBIT/TA	BVE/BVL	78.5	78.5	78.5
#17	EBITDA/I	RE/TA	EBITDA/IC	BVE/BVL	80.4	83.5	77.2
#18	EBITDA/I	RE/TA	EBITDA/OA	BVE/BVL	82.9	88.6	77.2

#19	EBITDA/I	RE/TA	EBIT/TA	FA/L	81.6	82.3	81.0
#20	EBITDA/I	RE/TA	EBIT/TA	FA/CL	78.5	77.2	79.7
#21	EBIT/I	ANI/ATA	EBIT/TA	BVE/BVL	79.7	77.2	82.3
#22	EBIT/I	AEBIT/AOA	EBIT/TA	BVE/BVL	79.1	78.5	79.7
#23	EBIT/I	RE/TA	EBITDA/IC	BVE/BVL	84.2	89.9	78.5
#24	EBIT/I	RE/TA	EBITDA/OA	BVE/BVL	84.2	89.9	78.5
Model C	EBIT/I	RE/TA	EBIT/TA	FA/L	82.3	83.5	81.0
#26	EBIT/I	RE/TA	EBIT/TA	FA/CL	80.4	79.7	81.0
#27	WC/TA	ANI/ATA	EBITDA/IC	BVE/BVL	78.5	75.9	81.0
#28	WC/TA	ANI/ATA	EBITDA/OA	BVE/BVL	81.0	86.1	75.9
Model B	WC/TA	ANI/ATA	EBIT/TA	FA/L	84.2	86.1	82.3
#30	WC/TA	ANI/ATA	EBIT/TA	FA/CL	79.7	78.5	81.0
#31	WC/TA	AEBIT/AOA	EBITDA/IC	BVE/BVL	77.2	81.0	73.4
#32	WC/TA	AEBIT/AOA	EBITDA/OA	BVE/BVL	81.0	87.3	74.7
#33	WC/TA	AEBIT/AOA	EBIT/TA	FA/L	82.9	84.8	81.0
#34	WC/TA	AEBIT/AOA	EBIT/TA	FA/CL	79.7	78.5	81.0
#35	WC/TA	RE/TA	EBITDA/IC	FA/L	75.9	82.3	69.6
#36	WC/TA	RE/TA	EBITDA/IC	FA/CL	78.5	83.5	73.4
Model A	WC/TA	RE/TA	EBITDA/OA	FA/L	86.1	93.7	78.5
#38	WC/TA	RE/TA	EBITDA/OA	FA/CL	83.5	87.3	79.7

Appendix 5 – Prediction accuracy changing three ratios

Model	**Var 1**	**Var 2**	**Var 3**	**Var 4**	**Overall%**	**B%**	**NB%**
Altman	WC/TA	RE/TA	EBIT/TA	BVE/BVL	77.8	74.7	81.0
#39	WC/TA	ANI/ATA	EBITDA/IC	FA/L	75.9	83.5	68.4
#40	WC/TA	ANI/ATA	EBITDA/IC	FA/CL	78.5	78.5	78.5
#41	WC/TA	ANI/ATA	EBITDA/OA	FA/L	81.6	91.1	72.2
#42	WC/TA	ANI/ATA	EBITDA/OA	FA/CL	79.7	84.8	74.7
#43	WC/TA	AEBIT/AOA	EBITDA/IC	FA/L	75.9	84.8	67.1
#44	WC/TA	AEBIT/AOA	EBITDA/IC	FA/CL	79.7	83.5	75.9
#45	WC/TA	AEBIT/AOA	EBITDA/OA	FA/L	83.5	93.7	73.4
#46	WC/TA	AEBIT/AOA	EBITDA/OA	FA/CL	79.1	84.8	73.4
#47	EBITDA/I	RE/TA	EBITDA/IC	FA/L	75.3	81.0	69.6
#48	EBITDA/I	RE/TA	EBITDA/IC	FA/CL	79.1	83.5	74.7
#49	EBITDA/I	RE/TA	EBITDA/OA	FA/L	83.5	91.1	75.9
#50	EBITDA/I	RE/TA	EBITDA/OA	FA/CL	82.3	87.3	77.2
#51	EBIT/I	RE/TA	EBITDA/IC	FA/L	77.2	83.5	70.9
#52	EBIT/I	RE/TA	EBITDA/IC	FA/CL	80.4	84.8	75.9
#53	EBIT/I	RE/TA	EBITDA/OA	FA/L	84.8	92.4	77.2
#54	EBIT/I	RE/TA	EBITDA/OA	FA/CL	84.2	91.1	77.2
#55	EBITDA/I	ANI/ATA	EBIT/TA	FA/L	82.3	83.5	81.0
#56	EBITDA/I	ANI/ATA	EBIT/TA	FA/CL	79.1	77.2	81.0
#57	EBITDA/I	AEBIT/AOA	EBIT/TA	FA/L	82.3	83.5	81.0
#58	EBITDA/I	AEBIT/AOA	EBIT/TA	FA/CL	79.1	77.2	81.0
#59	EBIT/I	ANI/ATA	EBIT/TA	FA/L	82.3	83.5	81.0
#60	EBIT/I	ANI/ATA	EBIT/TA	FA/CL	79.7	78.5	81.0
#61	EBIT/I	AEBIT/AOA	EBIT/TA	FA/L	82.3	83.5	81.0
#62	EBIT/I	AEBIT/AOA	EBIT/TA	FA/CL	79.1	78.5	79.7
#63	EBITDA/I	ANI/ATA	EBITDA/IC	BVE/BVL	79.7	81.0	78.5
#64	EBITDA/I	ANI/ATA	EBITDA/OA	BVE/BVL	81.6	87.3	75.9
#65	EBITDA/I	AEBIT/AOA	EBITDA/IC	BVE/BVL	75.3	86.1	64.6
#66	EBITDA/I	AEBIT/AOA	EBITDA/OA	BVE/BVL	81.6	88.6	74.7

#67	EBIT/I	ANI/ATA	EBITDA/IC	BVE/BVL	81.6	84.8	78.5
#68	EBIT/I	ANI/ATA	EBITDA/OA	BVE/BVL	83.5	89.9	77.2
#69	EBIT/I	AEBIT/AOA	EBITDA/IC	BVE/BVL	77.2	87.3	67.1
#70	EBIT/I	AEBIT/AOA	EBITDA/OA	BVE/BVL	81.6	88.6	74.7

Appendix 6 – Prediction accuracy changing four ratios

Model	Var 1	Var 2	Var 3	Var 4	Overall%	B%	NB%
Altman	WC/TA	RE/TA	EBIT/TA	BVE/BVL	77.8	74.7	81.0
#71	EBITDA/I	ANI/ATA	EBITDA/IC	FA/L	75.3	83.5	67.1
#72	EBITDA/I	ANI/ATA	EBITDA/IC	FA/CL	77.8	78.5	77.2
#73	EBITDA/I	ANI/ATA	EBITDA/OA	FA/L	81.0	91.1	70.9
#74	EBITDA/I	ANI/ATA	EBITDA/OA	FA/CL	79.7	84.8	74.7
#75	EBITDA/I	AEBIT/AOA	EBITDA/IC	FA/L	75.9	88.6	63.3
#76	EBITDA/I	AEBIT/AOA	EBITDA/IC	FA/CL	77.8	86.1	69.6
#77	EBITDA/I	AEBIT/AOA	EBITDA/OA	FA/L	82.9	93.7	72.2
#78	EBITDA/I	AEBIT/AOA	EBITDA/OA	FA/CL	81.0	86.1	75.9
#79	EBIT/I	ANI/ATA	EBITDA/IC	FA/L	77.8	87.3	68.4
#80	EBIT/I	ANI/ATA	EBITDA/IC	FA/CL	77.2	81.0	73.4
#81	EBIT/I	ANI/ATA	EBITDA/OA	FA/L	81.6	91.1	72.2
#82	EBIT/I	ANI/ATA	EBITDA/OA	FA/CL	81.6	88.6	74.7
#83	EBIT/I	AEBIT/AOA	EBITDA/IC	FA/L	75.9	88.6	63.3
#84	EBIT/I	AEBIT/AOA	EBITDA/IC	FA/CL	78.5	87.3	69.6
#85	EBIT/I	AEBIT/AOA	EBITDA/OA	FA/L	83.5	93.7	73.4
#86	EBIT/I	AEBIT/AOA	EBITDA/OA	FA/CL	79.1	86.1	72.2

Appendix 7 – Test of significance of select models

Model	Wilks' Lambda	Chi-square	df	Sig.
Altman	0.612	75.729	4	0.000
Model A	0.689	57.468	4	0.000
Model B	0.603	77.975	4	0.000
Model C	0.592	80.753	4	0.000

Appendix 8 – Standardized Canonical Discriminant Function Coefficients

Altman		Model A		Model B		Model C	
WC/TA	-0.297	WC/TA	-0.222	WC/TA	-0.118	EBIT/I	0.175
RE/TA	0.412	RE/TA	0.396	ANI/ATA	0.067	RE/TA	0.179
EBIT/TA	0.910	EBITDA/OA	0.721	EBIT/TA	0.849	EBIT/TA	0.764
BVE/BVL	0.128	FA/L	0.311	FA/L	0.425	FA/L	0.281

Appendix 9 – Eigenvalues

Model	Eigenvalue	Canonical correlation
Altman	0.635	0.623
Model A	0.452	0.558
Model B	0.659	0.630
Model C	0.689	0.639

Appendix 10 – McNemar's Test

	Model Z´´ hit	Model Z´´ miss	Total
Model A hit	120	16	136
Model A miss	3	19	22
Total	123	35	158

P (two tailed)	0.00443

	Model Z´´ hit	Model Z´´ miss	Total
Model B hit	122	11	136
Model B miss	1	24	22
Total	123	35	158

P (two tailed)	0.00635

	Model Z´´ hit	Model Z´´ miss	Total
Model C hit	121	9	136
Model C miss	2	26	22
Total	123	35	158

P (two tailed)	0.06542

Appendix 11 – Test of equality of group means

Altman Z-score model					
	B(mean)	NB(mean)	Wilk's Lambda	F	Sig.
WC/TA	0.007	0.207	0.903	16.679	0.000
RE/TA	-0.046	0.254	0.840	29.606	0.000
EBIT/TA	-0.096	0.094	0.647	85.253	0.000
BVE/BVL	0.183	0.564	0.902	16.897	0.000

Model A					
	B(mean)	NB(mean)	Wilk's Lambda	F	Sig.
WC/TA	0.007	0.207	0.903	16.679	0.000
RE/TA	-0.046	0.254	0.840	29.606	0.000
EBITDA/OA	-0.074	0.254	0.736	55.924	0.000
FA/L	0.163	0.418	0.830	31.947	0.000

Model B					
	B(mean)	NB(mean)	Wilk's Lambda	F	Sig.
WC/TA	0.007	0.207	0.903	16.679	0.000
ANI/ATA	-0.059	0.048	0.828	32.335	0.000
EBIT/TA	-0.096	0.094	0.647	85.253	0.000
FA/L	0.163	0.418	0.830	31.947	0.000

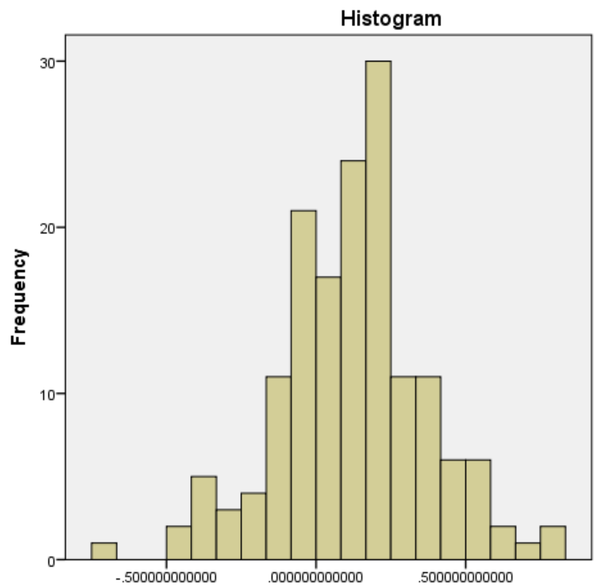
Model C					
	B(mean)	NB(mean)	Wilk's Lambda	F	Sig.
EBIT/I	-0.050	0.130	0.893	18.603	0.000
RE/TA	-0.046	0.254	0.840	29.606	0.000
EBIT/TA	-0.096	0.094	0.647	85.253	0.000
FA/L	0.163	0.418	0.830	31.947	0.000

Appendix 12 – Shapiro-Wilk test for normality

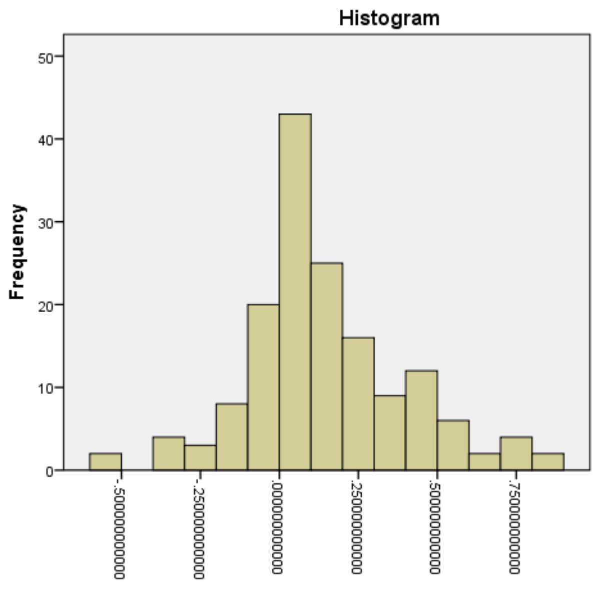
Variable	**Statistic**	**Df**	**Sig.**
WC/TA	0.988	158	0.208
RE/TA	0.964	158	0.000
EBIT/TA	0.990	158	0.353
BVE/BVL	0.829	158	0.000
EBIT/I	0.455	158	0.000
ANI/ATA	0.977	158	0.012
EBITDA/OA	0.990	158	0.322
FA/L	0.831	158	0.000

Appendix 13 – Histograms of frequency of observations (ratios)

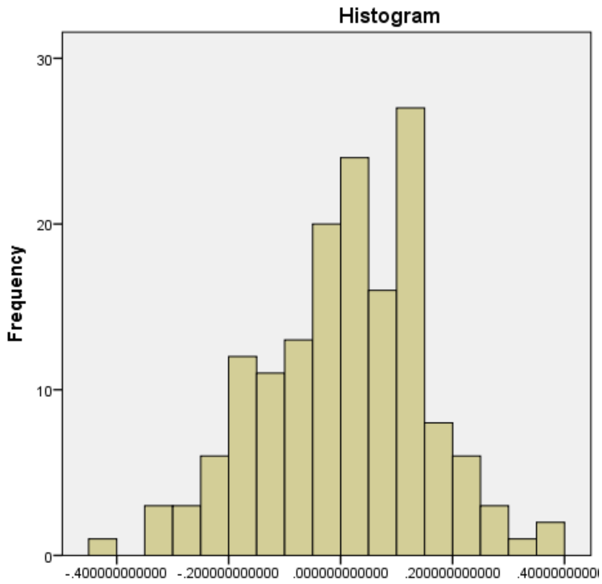
WC/TA



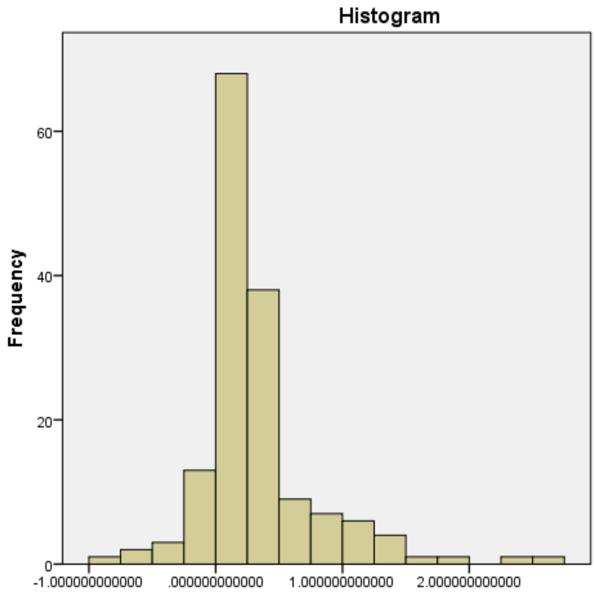
RE/TA



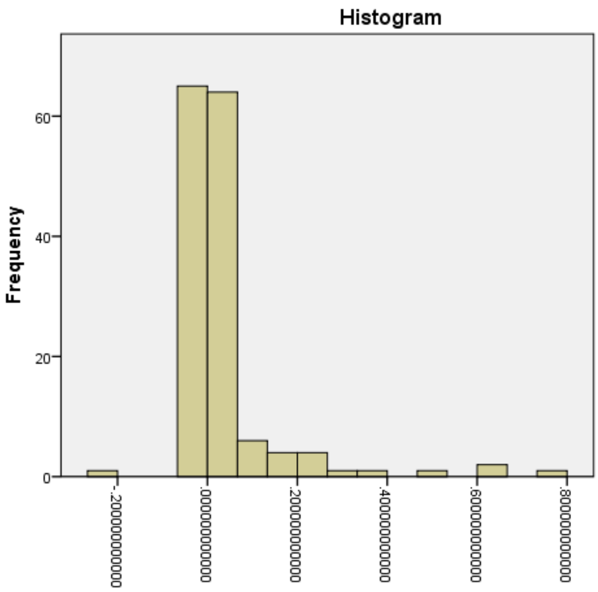
EBIT/TA



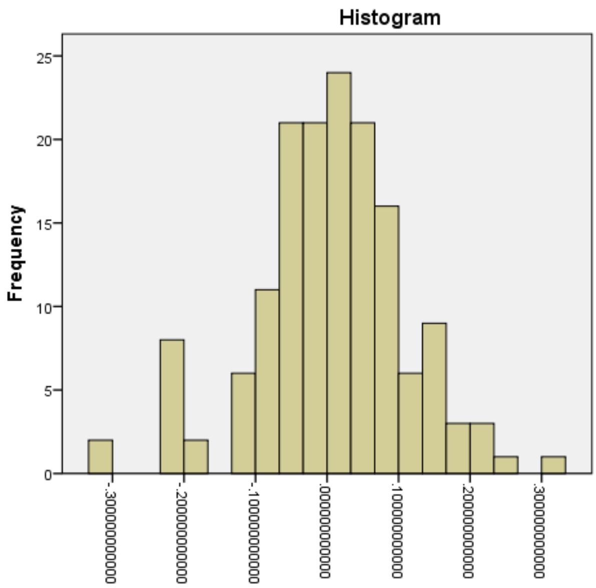
BVE/BVL



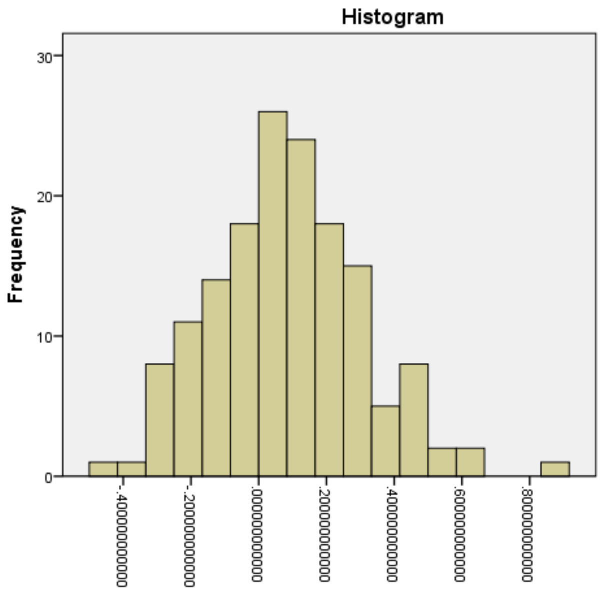
EBIT/I



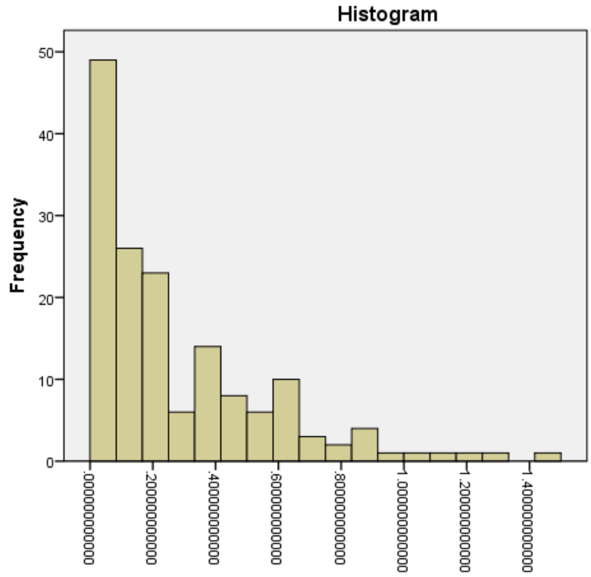
ANI/ATA



EBITDA/OA

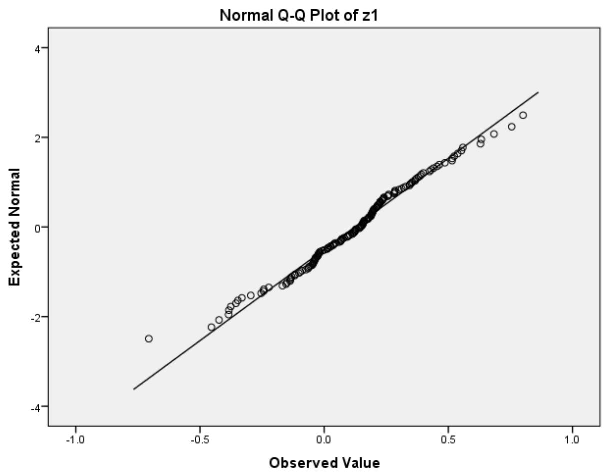


FA/L

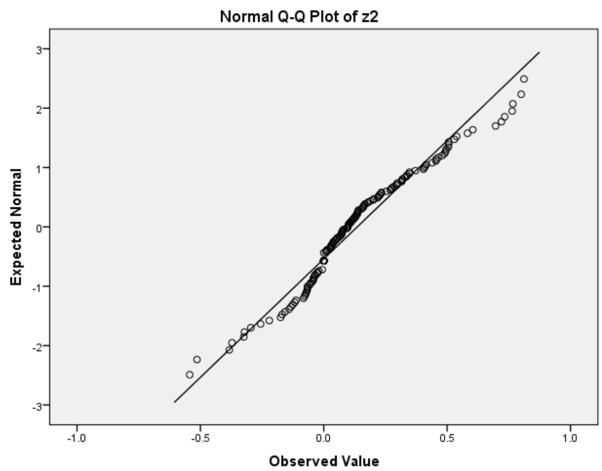


Appendix 14 – Normal Q-Q plots (ratios)

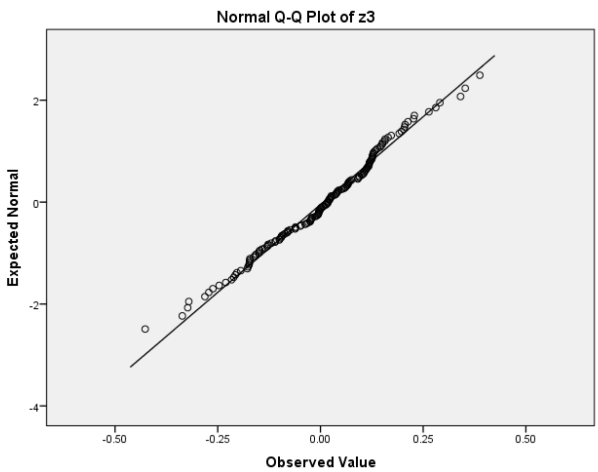
WC/TA



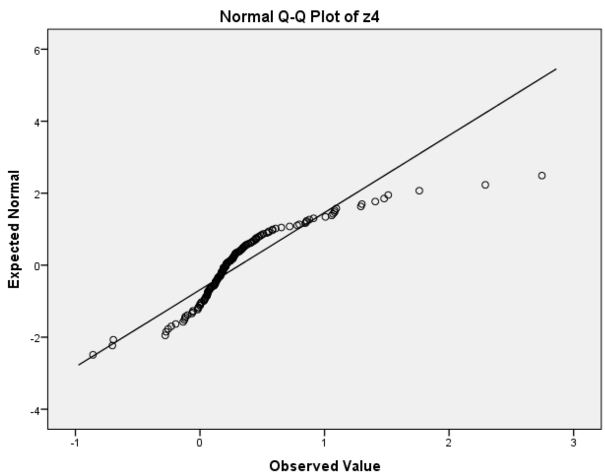
RE/TA



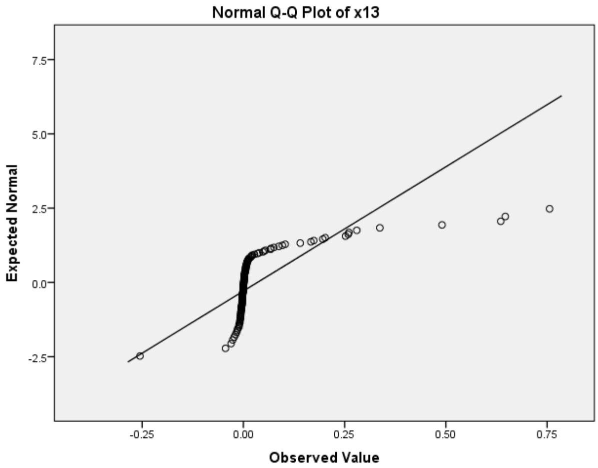
EBIT/TA



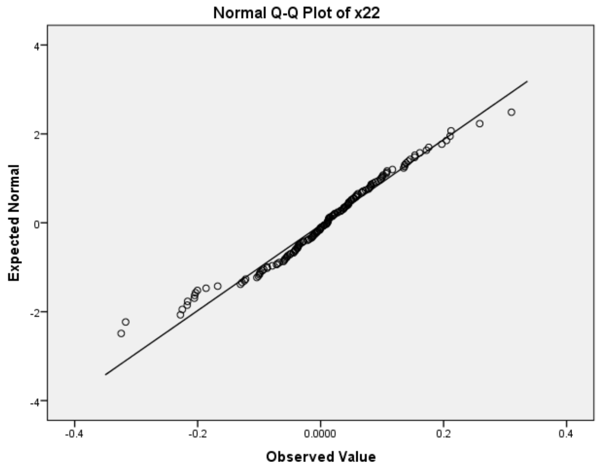
BVE/BVL



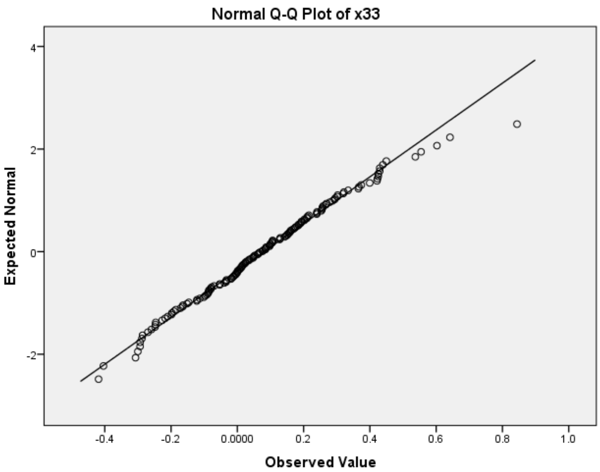
EBIT/I



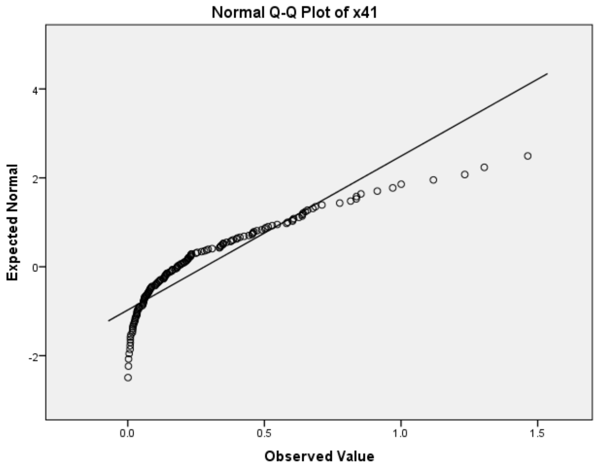
ANI/ATA



EBITDA/OA

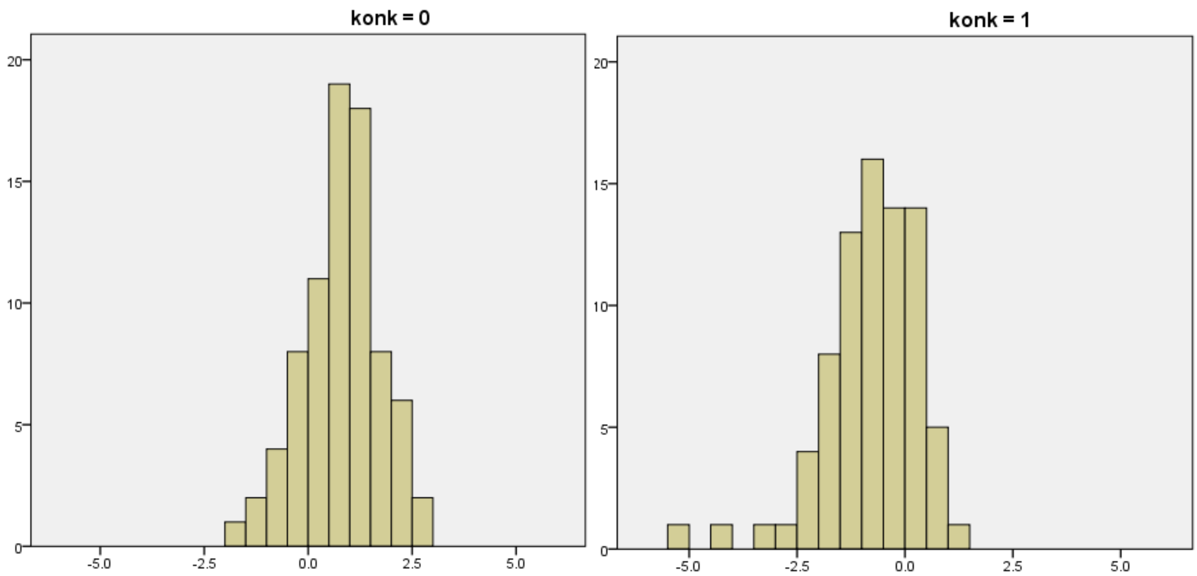


FA/L

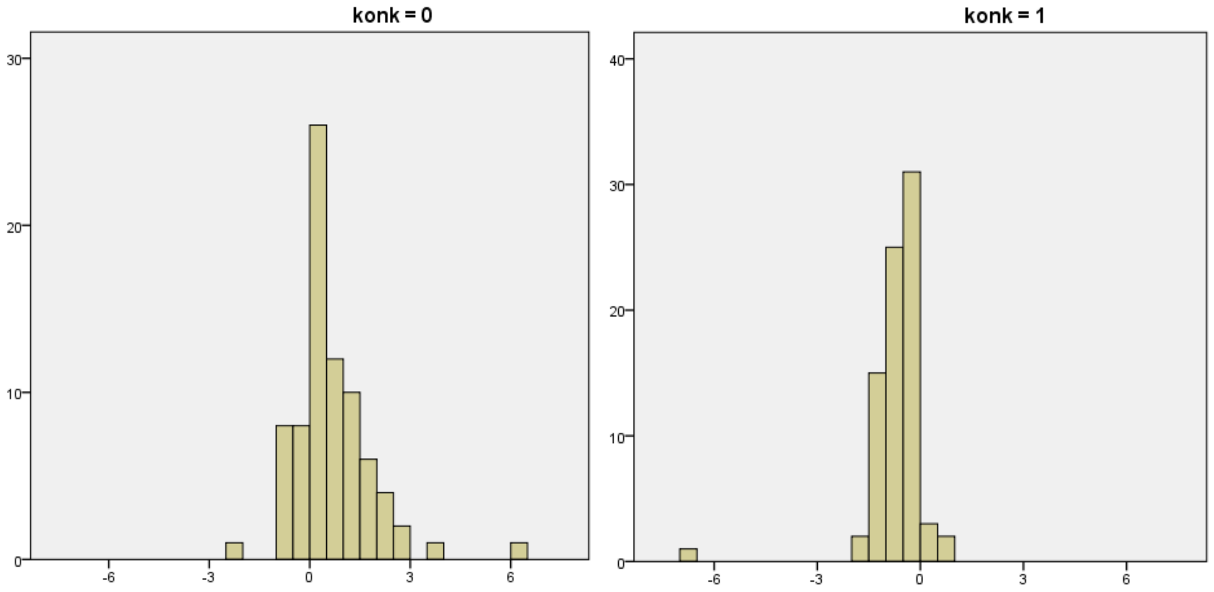


Appendix 15 – Histograms of frequency of observations (groups)

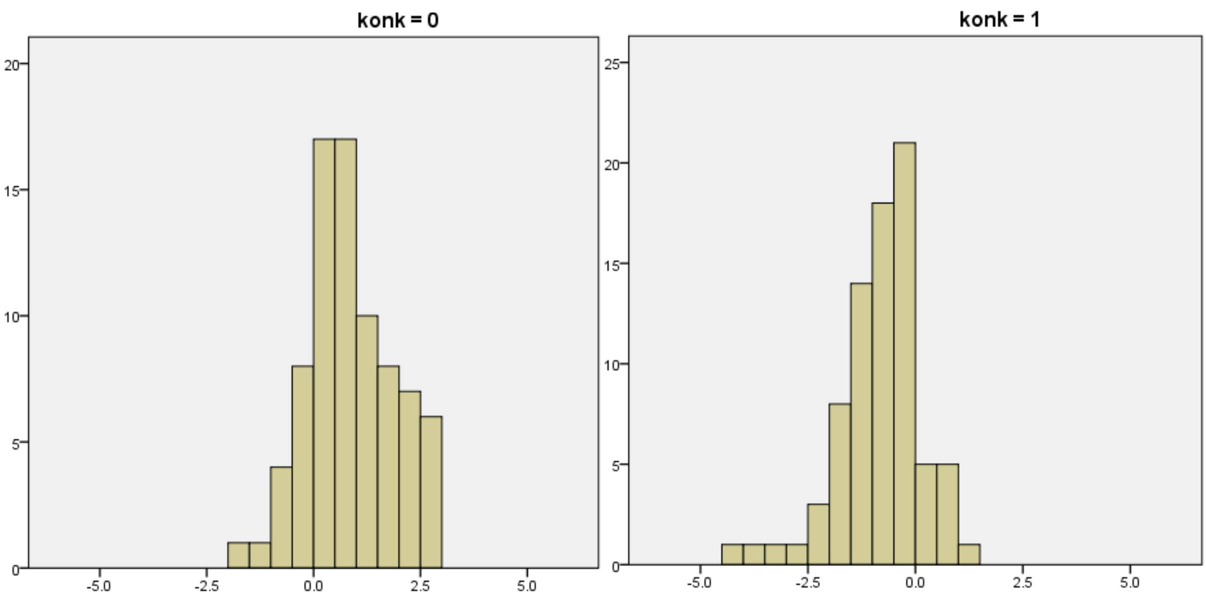
Altman's Z´´-score model



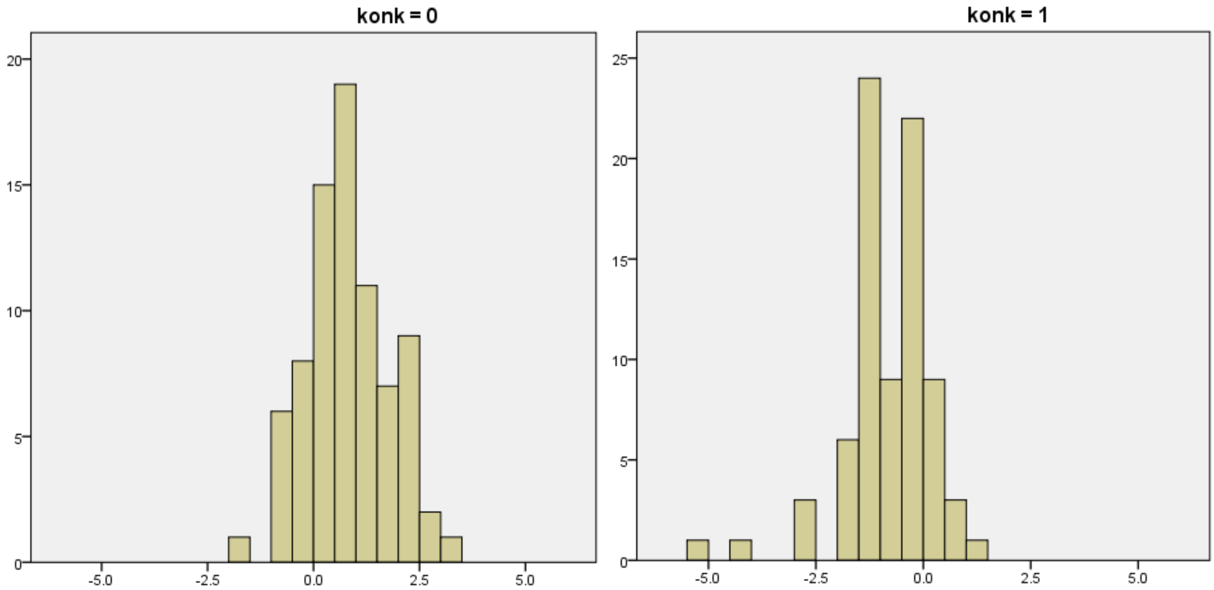
Model A



Model B



Model C



Appendix 16 – Box's M test

		Altman	Model A	Model B	Model C
Box's M		131.729	125.450	69.009	77.799
F	Approx.	12.809	12.198	6.710	7.565
	df1	10	10	10	10
	df2	116347.410	116347.410	116347.410	116347.410
	Sig.	0.000	0.000	0.000	0.000

Appendix 17 – Log determinants

	Altman		Model A		Model B		Model C	
	Rank	Log det	Rank	Log det	Rank	Log det	Rank	Log det
Non-bankrupt	4	-10.821	4	-10.373	4	-14.350	4	-11.198
Bankrupt	4	-12.405	4	-12.483	4	-14.356	4	-13.333
Pooled within group	4	-10.769	4	-10.624	4	-13.912	4	-11.767

Appendix 18 – Pooled Within-Groups matrices

Altman		**WC/TA**	**RE/TA**	**EBIT/TA**	**BVE/BVL**
Covariance	**WC/TA**	0.095	0.068	0.017	0.086
	RE/TA	0.068	0.120	0.013	0.112
	EBIT/TA	0.017	0.013	0.017	0.017
	BVE/BVL	0.086	0.112	0.017	0.340
Correlation	**WC/TA**	1.000	0.638	0.421	0.479
	RE/TA	0.638	1.000	0.279	0.553
	EBIT/TA	0.421	0.279	1.000	0.219
	BVE/BVL	0.479	0.553	0.219	1.000

Model A		**WC/TA**	**RE/TA**	**EBITDA/OA**	**FA/L**
Covariance	**WC/TA**	0.095	0.068	0.042	0.028
	RE/TA	0.068	0.120	0.037	0.036
	EBITDA/OA	0.042	0.037	0.076	0.031
	FA/L	0.028	0.036	0.031	0.081
Correlation	**WC/TA**	1.000	0.638	0.493	0.321
	RE/TA	0.638	1.000	0.388	0.364
	EBITDA/OA	0.493	0.388	1.000	0.401
	FA/L	0.321	0.364	0.401	1.000

Model B		**WC/TA**	**ANI/ATA**	**EBIT/TA**	**FA/L**
Covariance	**WC/TA**	0.095	0.014	0.017	0.028
	ANI/ATA	0.014	0.014	0.008	0.008
	EBIT/TA	0.017	0.008	0.017	0.007
	FA/L	0.028	0.008	0.007	0.081
Correlation	**WC/TA**	1.000	0.394	0.421	0.321
	ANI/ATA	0.394	1.000	0.514	0.244
	EBIT/TA	0.421	0.514	1.000	0.181
	FA/L	0.321	0.244	0.181	1.000

Model C		**EBIT/I**	**RE/TA**	**EBIT/TA**	**FA/L**
Covariance	**EBIT/I**	0.069	0.016	0.005	0.026
	RE/TA	0.016	0.120	0.013	0.036
	EBIT/TA	0.005	0.013	0.017	0.007
	FA/L	0.026	0.036	0.007	0.081
Correlation	**EBITDA/I**	1.000	0.175	0.146	0.348
	RE/TA	0.175	1.000	0.279	0.364
	EBIT/TA	0.146	0.279	1.000	0.181
	FA/L	0.348	0.364	0.181	1.000