

ABSOLUTE & RELATIVE CREDIT QUALITY ASSESSMENT

Svante Dieden Sandell & Mattias Karlsson

June 2016

Abstract

The lack of availability and relevance of both credit ratings and traded market instruments, forces financial institutions to find alternative ways to validate the credit qualities of their counterparties. To address this issue, existing bankruptcy prediction models are evaluated and re-estimated. Furthermore a new model is constructed that outperforms the previous models in terms of default classification. By adjusting for the rarity of defaults and the utilised sampling techniques, the output of the constructed model becomes more accurate and less biased than previous models. The model is also validated to be rank consistent with US and Nordic S&P ratings as well as with spreads of Credit Default Swaps on the US market.

Keywords: Credit Quality Assessment, Default & Bankruptcy Prediction, Altman Z-score, Ohlson O-score, Logistic Regression, Rare-Event Bias Correction.

Acknowledgements

The research included in this thesis could not have been performed if not for the assistance, patience, and support of many individuals. We would first and foremost like to extend our gratitude to our supervisor Associate Professor Magnus Wiktorsson at the department of Mathematical Statistics, Lund University, who has been extremely helpful and has provided us with stimulating discussions. We would also like to extend our sincere thanks to Babak Soltani, Ingvar Matsson, Klas Andreasen, Pia Hellborg, Ingrid Harbo, and the rest of the Internal Audit team at Swedbank who provided us with the opportunity to conduct research for them, and who gave access to the trading floor, equipment and research facilities utilised in this thesis. Without their support it would not have been possible to conduct this research. Finally we would like to thank our families for their continuous support.

Svante Dieden Sandell & Mattias Karlsson
Lund, 1st of June 2016

Contents

1	Introduction	1
1.1	Background	1
1.2	Relevant Literature	2
1.3	Problem Discussion	3
1.4	Problem Formulation	4
1.5	Thesis Outline	5
2	Theory and Concepts	6
2.1	Statistical Theory	6
2.2	Financial Theory	18
2.3	Earlier Models	21
2.4	Model Evaluation Theory	24
3	Data	28
3.1	Credit Event Sample	28
3.2	Non-Credit Event Sample	30
3.3	Final Adjustments to the Samples	30
3.4	Ratios	32
3.5	Estimation & Validation Sets	33
3.6	Market Data	33
4	Method	35
4.1	Model Building	35
4.2	Earlier Models	40
4.3	Model Evaluation	40
5	Results	41
5.1	Model Building	41
5.2	Earlier Models	54
5.3	Model Evaluation	58
6	Discussion	65
6.1	The 5-Factor Model	65
6.2	The Performance of the Models	68

6.3	Model Building	71
6.4	The Model in Practice	74
7	Suggestions for Further Research	76
8	Conclusion	77
	Appendices	79
A	Construction of Non-Credit Event Sample	80
B	Bloomberg Variables & Definitions	82
C	Credit Rating Data & CDS Data	85
D	Flexible Data Set - Pseudo Code	87
E	Stepwise Inclusion/Exclusion - Pseudo Code	88
F	Descriptive Statistics	90
G	Covariate Family Assignment	92
H	Correlation Matrix	94
I	Representative Visual Examples	96
J	Histogram for Final Ratios	99
K	Re-Estimated Altman & Ohlson	100
	Bibliography	102

1

Introduction

"Begin at the beginning," the King said, gravely, "and go on till you come to an end; then stop."

– Lewis Carroll

1.1 Background

1.1.1 Credit Quality

Since the start of the financial services industry, credit quality assessment has constituted an integral part of financial institutions' core business. In essence, credit quality is a measure of a debtors ability to cover its financial obligations. If a debtor can not meet its obligations, i.e. experience a *credit event*, it will result in losses for its counterparties.

In the aftermath of the financial crisis, credit quality assessment has become increasingly important. A plethora of regulations has been introduced that forces financial institutions to hold additional capital to offset counterparty risk. Restrictions on the amount of capital available for investments decrease the potential profitability of financial institutions. By having precise credit quality assessments, the amount of capital held and the fees assigned to contracts can be made as accurate as possible. It is a game where one balances temporary profitability against severe unexpected losses while staying within the regulatory boundaries. As is common in the financial industry, having access to the best information is key to long-term profitability.

There are three main types of corporate credit quality assessment sources addressed in this thesis, (1) Third party credit ratings; (2) Traded spreads on CDS-contracts or bonds; and (3) Credit and bankruptcy prediction models. This thesis focuses mainly on the third type, even though the first two are the more commonly known.

1.1.2 Credit Ratings

The need for credit quality assessments has given rise to the creation of large credit rating agencies, such as Moody's, Standard and Poor's (S&P), and Fitch Group. The primary purpose of these agencies is to set credit grades, commonly known as ratings, on companies, bonds and other financial instruments. A higher rating of a company is viewed as positive as it lowers the company's expected cost of new capital. Ratings are also observed closely by the rated companies' counterparties as a safer company has, all else being equal, a lower expected loss given default, which intuitively should require less capital to be held.

1.1.3 Market Implied Credit Quality

The credit qualities of companies can, fortunately or unfortunately, also be implied by assets traded on the market, such as Credit Default Swaps (CDS) or corporate bonds. While credit rating agencies can take long time between their credit rating updates, the CDS and bond prices are in almost all cases updated more frequently as they are traded on the market. The soaring CDS prices seen during the financial crisis raises the question whether the CDS contracts are really limited to capturing the company specific default risk. Some claim that other factors, such as the liquidity of the contracts play an important part in the pricing.

1.1.4 Model Implied Credit Quality

A model implied credit quality assessment attempts to independently of rating agencies, accurately and objectively assess credit quality in and out of time of crisis. There has been a considerable amount of academic research developed around bankruptcy prediction and credit scoring models. The output of these models can be used for objective credit quality assessment. Two famous models are Altman's Z-score and Ohlson's O-Score, both of which are used by practitioners. Many of the academic models, including Ohlson's O-score, yield a nominal, rather than ordinal, value interpretable as the probability of bankruptcy.

1.2 Relevant Literature

One of the classic works in the area of bankruptcy prediction was conducted by Beaver (1966). He performed univariate analysis for a number of bankruptcy predictors and set the stage for the multivariate analyses that followed. Beaver (1966) found several predictors that can discriminate between matched samples of failed and non-failed firms.

Altman (1968) is the first one using multivariate analysis of ratios to predict bankruptcies. In his well-known paper he uses data from American manufacturing companies ranging from 1946 to 1965. Altman shows that by using a combination of observed accounting ratios it is possible to predict corporate bankruptcy. Altman's research was groundbreaking at the time and kindled an interest for similar analyses.

A logistic regression approach is undertaken by Ohlson (1980) for prediction of bankruptcies within American industrial companies using data ranging from 1970 to 1976. His model uses variables similar to that of Altman's Z-score model but his approach mitigates some of the critique directed at the statistical technique utilised by Altman (1968).

With a non-financial background and with applications to social sciences in mind, King and Zeng (2001) exhibit and provide solutions to issues surrounding sampling and rare-event bias introduced by sloppy application of sampling techniques. In their paper they also criticise the lack of consideration for these problems seen from applied statisticians. Using the methodology proposed by King and Zeng (2001) is supposed to give more accurate probability estimates when conducting logistic regression for rare-event prediction.

1.3 Problem Discussion

Main issues with credit quality assessment concerns the availability and reliability of data. Take credit ratings from the third party rating agencies as an example. Are there ratings available for the company of interest? Are these ratings, if they exist, truly unbiased? To answer the first question, consider the Nordic market, where only a few, of the biggest companies, have ratings provided by the major agencies. The applicability of credit ratings on the Nordic market is therefore limited. One reason giving doubt to the second question is that companies can pay for earlier credit rating updates. Intuition suggests that such payments are only done when a higher rating is likely to be given. This gives rise to a possible bias towards too high ratings as well as a lagged introduction of lower ratings.

An alternative approach to the credit quality assessment conundrum presents itself by consideration of CDS-data observable on the market. But the previously discussed sparsity issue, present for the rating data, is an arguably even greater issue for the CDS data. Although the CDS-contracts are updated more frequently and constructed to capture the true credit event probabilities, the CDS-spreads suffer from liquidity issues, are impacted by herd behaviour and are by their contractual nature driven by supply and demand.

As a third option one can instead rely on famous models such as Altman's Z-score and Ohlson's O-score for implied credit quality, but it has been

decades since Altman and Ohlson constructed their respective models. In the meantime, the world, and most companies in it, have changed. It is therefore reasonable to question if their models are still applicable today, and if so, are they accurate? Besides, as companies and the financial industry have evolved it is likely that other ratios, than those originally considered by Altman and Ohlson, are better suited for credit quality assessment today.

Just as much as the fundamentals of companies may have changed, statistical theory has also evolved. Consideration of relatively recent advances in regression analysis enables more accurate estimates of true credit event probabilities. For example, although Ohlson's model has probabilities as output, these are severely biased and not taking the population wide-probability of default into account. They are instead calibrated to the sample probability of default. The expected average output from the models should be close to the true population-wide default rate, but this is not the case as they have been calibrated on a biased subset of the population. By applying the techniques suggested by King and Zeng (2001) it should be possible to come closer to the true probabilities.

For a model to be applicable for credit quality assessment, not only the absolute probability estimates are of interest. If one intends to use the model for companies that have no ratings or CDS-spreads, and one additionally wants to assess a proxy rating or CDS-spread to these companies, then at least some rank consistencies between the model and these data types are desired. If rank consistency exists then an un-rated company could have its credit quality assessed without the need to consult the credit rating agencies. Consider an example where three companies are present, two already have ratings and a proxy rating is requested for the third. It is then possible to, by use of a model, for which rank consistency is relatively strong, rate the third company by ordinal comparison. For example, say that the first and second company have a rating and model score pair of (B, 3 %) and (AA+, 0.5 %) respectively. Then the third company can be given a proxy rating depending on the model score. For a model score between 0 % and 0.5 % a rating of at least AA+ is implied. if the model score is between 0.5 % and 3 % then the corresponding implied rating is between B and AA+ and finally a model score greater than 3 % would imply a lower rating than B. The same relative ranking procedure would also be applicable to CDS-spreads if the model's output is concluded to be relatively rank consistent to CDS-spreads.

1.4 Problem Formulation

The questions addressed in this thesis are:

- How do Altman's Z-score and Ohlson's O-score perform on a more recent data set?

- Will a re-estimation of Altman's Z-score and Ohlson's O-score, to better suit recent market conditions, improve their performances?
- Is it possible to construct a logistic regression model which performs superior to Altman's and Ohlson's models for prediction of credit events?
- Can the constructed logistic regression model's output probabilities be made more realistic?
- Is the constructed logistic regression model rank consistent with the credit qualities implied by S&P and CDS-spreads?

1.5 Thesis Outline

The outline for the rest of this thesis is as follows:

Chapter 2: Statistical and financial theory will be introduced, along with concepts that are of importance for this report.

Chapter 3: The common denominator for all the models considered in this report is the need for credit event and non-credit event companies. For each company the models require data from exactly one annual statement. The gathering process along with a description of the retrieved data is described in detail in this chapter.

Chapter 4: In this chapter the following issues are resolved, (1) How a logistic regression model is built from the data described in Chapter 3; (2) How different models are evaluated against each other; (3) How model output is tested for rank consistency with CDS and Rating data.

Chapter 5: This chapter simply presents results obtained by following the methodology outlined in Chapter 4.

Chapter 6: In this chapter, (1) The resulting model from the model building stage is presented and discussed in detail; (2) All models' performances are discussed and the performance is related to CDS and Rating data; (3) Emerging issues from the model building stage are discussed, along with other important clarifications; (4) It is presented how the resulting model can be used in practice.

Chapter 7: Interesting suggestions for future research are presented here.

Chapter 8: The final chapter summarises and concludes the thesis.

2

Theory and Concepts

Everything should be made as simple as possible, but not simpler.
– Albert Einstein

2.1 Statistical Theory

2.1.1 Logistic Regression

This thesis concerns explanation and prediction of credit events through in advance observable data. Since the occurrence of a credit event is a dichotomous event, a linear regression would fail to capture the dynamics of the response-variable. Instead a more general class of regression tools is utilised in this report, namely generalised linear models (GLM), see for instance Dobson and Barnett (2007) for an introduction to the subject. Logistic regression is a special case within the GLM-family and it is useful when the response is a binary categorical variable.

The goal of logistic regression is to explain the relation between the k explanatory variables in the column vector $\mathbf{X}_i = \{1, X_{i,1}, \dots, X_{i,k}\}^T$ and the outcome variables Y_i , for $i = 1, \dots, n$. In logistic regression Y_i is a binary response s.t. $Y_i \sim \text{Bernoulli}(p_i)$. Y_i thus takes on the value 1 w.p. p_i and takes on the value 0 w.p. $1 - p_i$, that is $P(Y_i = y_i) = p_i^{y_i}(1 - p_i)^{1-y_i}$, for $y_i = 0, 1$. The value p_i is thought to follow an inverse logistic function of a $(k + 1) \times 1$ vector \mathbf{x}_i ,

$$p_i = \frac{1}{1 + e^{-\mathbf{x}_i^T \cdot \boldsymbol{\beta}}} \quad (2.1)$$

where all \mathbf{x}_i are, jointly independent, observations of \mathbf{X}_i , for $i = 1, \dots, n$. Furthermore, $\boldsymbol{\beta}$ is a $(k + 1) \times 1$ vector, with an intercept in the first row. The objective is to calibrate the vector $\boldsymbol{\beta}$ so that for each new set of explanatory variables, \mathbf{x}_i the model gives a probability that the corresponding response is a successful Bernoulli event, i.e. in this thesis a credit event. (Agresti, 2007; King & Zeng, 2001)

2.1.2 Maximum Likelihood

Consider a sample of n random vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, with joint distribution function, $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n|\theta)$ where $\theta \in \Theta$, and Θ is the parameter space. Define the likelihood function $L(\theta) = f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n|\theta)$ and observe that, given a sample $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, this is a function of θ . Furthermore if $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are mutually independent then $L(\theta) = f(\mathbf{x}_1|\theta) \cdot \dots \cdot f(\mathbf{x}_n|\theta)$, where $f(\mathbf{x}_i|\theta) = \int_{-\infty}^{\infty} f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n|\theta) d\mathbf{x}_1, \dots, d\mathbf{x}_{i-1}, d\mathbf{x}_{i+1}, d\mathbf{x}_n$. The maximum likelihood estimate (MLE) of θ is the value $\hat{\theta} = \max_{\Theta} L(\theta)$. Furthermore define $\log(L(\cdot))$ to be the log-likelihood function and note that since the logarithm function is monotone, the maximum of the log-likelihood function is obtained at the MLE $\hat{\theta}$, which is normally distributed under regularity conditions (“The Concise Encyclopedia of Statistics,” 2008). The *Score Function* is furthermore defined as the gradient of the log-likelihood function, and the (expected) *Fisher Information Matrix* $I(\theta)$ is defined as the variance of the *Score Function*, lastly the inverse $I(\theta)^{-1}$ is the asymptotic variance of the MLE. (Geyer, 2003) The consistent estimate $I(\hat{\theta})^{-1}$ is often used as a proxy when the true parameter θ is unknown.

2.1.3 Wald Test

In order to test if the MLE $\hat{\theta}$ differ significantly from zero it is possible to use the Wald test statistic. From asymptotic theory of MLE the difference between the estimated coefficient $\hat{\theta}$ and its corresponding true mean under H_0 (often set to $\theta = 0$) will be approximately normally distributed with mean θ . By subtracting the true mean from the MLE and dividing the result with the standard deviation of the MLE the result is a standard normal distribution if H_0 is true. By knowing the distribution a *p-value* is easily calculated.^a(Harrell, 2015)

2.1.4 Monte Carlo

Given that it is possible to generate an i.i.d. sequence X_1, X_2, \dots, X_n with common density f and given an integral of interest, $I = \int m(x)f(x)dx$. Then relying on the strong law of large numbers and the central limit theorem $\hat{I} = \frac{1}{n} \sum m(X_i)$ is a consistent, unbiased and asymptotically normal estimate of I . (Zamar, 2014)

If one can induce negative correlation within the sequence X_1, \dots, X_n the variance of the estimated integral \hat{I} will decrease. A crude way to do this is by using *antithetic variates*, where one, if possible, simply set X_{n+1}, \dots, X_{n+n} equal to $-X_1, \dots, -X_n$ to construct an additional n observations of the sequence. (Givens & Hoeting, 2013)

^aThis is the method used by MATLAB for logistic regression p-value calculations. For example the resulting value can be tested to the 5 % significance level by comparing the value obtained to ± 1.96 . If the resulting value is outside this level it is deemed significant.

2.1.5 Two-Sample t-Test

Equal Variance

A two-sample t-test can be used to determine if two populations' means differ significantly or not. The test involves testing the following hypotheses,

$$\begin{aligned} H_0 : \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i} &= \frac{1}{n_2} \sum_{j=1}^{n_2} X_{2,j} \\ H_1 : \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i} &\neq \frac{1}{n_2} \sum_{j=1}^{n_2} X_{2,j} \end{aligned} \quad (2.2)$$

The populations from which the samples are drawn should be normally distributed. The normality assumption should be tested for both samples independently. The standard deviations of the two populations should also be equal for the *equal variance two-sample t-test*. There is no requirement of equal size of the samples. The test statistic is calculated as follows,

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2.3)$$

where

$$s_{X_1 X_2} = \sqrt{\frac{(n_1 - 1) \cdot s_{X_1}^2 + (n_2 - 1) \cdot s_{X_2}^2}{n_1 + n_2 - 2}} \quad (2.4)$$

The null hypothesis should be rejected at significance level α if $|t| > t_{1-\alpha/2, v}$, where $t_{1-\alpha/2, v}$ is the critical value of the t distribution with v degrees of freedom, calculated as $v = n_1 + n_2 - 2$ when assuming equal variance. (“The Concise Encyclopedia of Statistics,” 2008)

Welch's t-Test (or Unequal Variance t-Test)

Welch's t-test is similar to the equal variance two-sample t-test, both determine if two populations have significantly different means or not. Ruxton (2006) argues that Welch's t-test should always be used instead of the equal variance t-test. Given the same hypotheses as above the t-statistic is calculated as follows,

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1, \bar{X}_2}} \quad (2.5)$$

where

$$s_{\bar{X}_1, \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (2.6)$$

For use in significance testing the distribution of the test statistic is approximated with an ordinary Student's t distribution with degrees of freedom calculated as,

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)} \quad (2.7)$$

Welch's t-test keeps the normality assumption of the populations from the equal variance t-test but differs with respect to the assumption of equal variance. (Welch, 1947)

2.1.6 Levene's Test & Brown Forsythe's Test

Levene's test (Levene, 1960) is used to test if two or more samples have significantly different variances or not. The Levene test for k samples is defined as,

$$\begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \dots = \sigma_k^2 \\ H_1 : \sigma_i^2 &\neq \sigma_j^2 \text{ for at least one pair } (i,j) \end{aligned} \quad (2.8)$$

The k samples have sample sizes n_1, \dots, n_k , s.t. $\sum_{i=1}^k n_i = N$ from k corresponding random variables X_1, \dots, X_k (which could be equally distributed). By X_{ij} the j 'th value, within sample i , is referred to. Levene's test statistic is defined as,

$$W = \frac{N - k}{k - 1} \cdot \frac{\sum_{i=1}^k n_i (\bar{Z}_i - \bar{Z}_{..})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2} \quad (2.9)$$

where Z_{ij} can have one of the three following definitions,

1. $Z_{ij} = |X_{ij} - \bar{X}_i|$, where \bar{X}_i is the mean of the i -th sample.
2. $Z_{ij} = |X_{ij} - \tilde{X}_i|$, where \tilde{X}_i is the median of the i -th sample.
3. $Z_{ij} = |X_{ij} - \tilde{X}'_i|$, where \tilde{X}'_i is the 10 % trimmed mean of the i -th sample.

\bar{Z}_i are the group means of the Z_{ij} , where $j = 1, \dots, n_i$, and $Z_{..}$ is the mean of all N values Z_{ij} .

Levene's original paper only proposed use of the first alternative definition of Z_{ij} . Brown and Forsythe (1974) extended Levene's test to use either the trimmed mean or the median. In this thesis Levene's test corresponds to the use of the mean and Brown Forsythe's test corresponds to the use of the median. The definition based on the median is usually recommended as the choice for non-normal data, as it is more robust. (NIST, 2012)

2.1.7 Kolmogorov-Smirnov Test

One Sample Test

The Kolmogorov-Smirnov (or K-S) test is a nonparametric test that determines if an i.i.d. sample X_1, \dots, X_n , drawn from an unknown distribution F , could have been drawn from a particular distribution F_0 at a given significance level. The hypothesis to test is as follows,

$$\begin{aligned} H_0 : F &= F_0 \\ H_1 : F &\neq F_0 \end{aligned} \tag{2.10}$$

Let $F(x) = P(X_1 \leq x)$, a cumulative distribution function (c.d.f.) of a true underlying distribution of the data. An empirical c.d.f. is furthermore defined as,

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \tag{2.11}$$

That counts the proportion of the sample points below level x . The strong law of large numbers implies that $F_n(x) \rightarrow F(x)$ a.s., and by the Gilvenko-Cantelli theorem the convergence is even uniform in x which gives an intuition for the Kolmogorov-Smirnov statistic as

$$\|F_n - F\|_\infty \equiv \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow{a.s.} 0 \tag{2.12}$$

Theorem 1. *Following the same notations as above. If $F(x)$ is continuous then the distribution of*

$$\sup_{t \in \mathbb{R}} |F_n(t) - F(t)|$$

does not depend on F .

Proof. See Panchenko (2006). □

Theorem 2. *Furthermore,*

$$\mathbb{P}(\sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \leq t) \rightarrow H(t) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 t}$$

where $H(t)$ is the c.d.f. of Kolmogorov-Smirnov distribution.

Proof. See Breiman (1968). □

The rule is to reject H_0 if $D_n > c$, where $D_n = \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$, and the threshold c depends on the significance level and can be found from the condition $\alpha = P(D_n \geq c | H_0)$. Under H_0 the distribution of D_n can be tabulated for each n and the threshold can be found. If n is large then the Kolmogorov-Smirnov distribution can be used to find c since $\alpha = P(D_n \geq c | H_0) \approx 1 - H(c)$. (Panchenko, 2006)

Two Sample Test

The Kolmogorov-Smirnov test for two samples works similar to the one-sample test. Suppose that for the first sample, m observations are drawn from a distribution with c.d.f. $F(x)$ and for the second sample n observations are drawn from a distribution with c.d.f. $G(x)$. The aim of the K-S two sample test is then to test,

$$\begin{aligned} H_0 &: F = G \\ H_1 &: F \neq G \end{aligned} \tag{2.13}$$

Denote the corresponding empirical distributions as $F_m(x)$ and $G_n(x)$. Then the test statistic is,

$$D_{mn} = \left(\frac{mn}{m+n}\right)^{1/2} \sup_x |F_m(x) - G_n(x)| \tag{2.14}$$

which satisfies theorem 1 and 2 above. And the rest is the same. (Panchenko, 2006)

2.1.8 Bonferroni Familywise Error Rate

When conducting multiple tests the risk of getting a False Positive is increasing in the number of tests performed. One can introduce the concept of Family Wise Error Rate (FWER) to correspond to the probability of making at least one type one error, i.e. a False Positive. To keep this FWER within control Bonferroni suggests to divide the significance level of each test by the number of times the test will be performed in total. From Boole's inequality it follows that this simple adjustment keeps the FWER below the predefined significance level. (Holm, 1979)

2.1.9 Correction for Choice Based Sampling

If a logistic regression model has discriminative power between, in this thesis, companies that have experienced a credit event (*credit event companies*) and companies that have not experienced credit events (*non-credit event companies*), then, in general, the credit event companies will be given larger probabilities, p , of experiencing credit events than the non-credit event companies. This is of course very natural. Another consequence in a rare event situation is that as events by definition are unlikely to occur then the estimated probabilities are rarely higher than 0.5. Thus, as a rule of thumb $p_{\text{Non-Credit Event Company}} < p_{\text{Credit Event Company}} < 0.5$. Furthermore, the asymptotic covariance matrix of the MLE $\hat{\beta}$ is, under regularity conditions^b, given by the inverse of the Information Matrix $I(\beta)^{-1}$, or a consistent estimate thereof, such as the inverse of the expected Information Matrix evaluated at the MLE $\hat{\beta}$, $\text{Var}(\hat{\beta}) = 1/(\sum_i p_i(1-p_i)x_i^T x_i)$, where $p_i = \frac{1}{1+e^{-x_i\hat{\beta}}}$.

Studying this function more closely, one notes that the constituent $p_i(1-p_i)$ where $p_i \in [0, 1]$ is maximised for $p_i = 0.5$ and the function is unimodal.

Conclusively, relying on the rule of thumb and the interpretation of the impact of probability estimates on the covariance matrix, it is noted that including additional credit event companies, instead of non-credit event companies, in the data sample contributes more to reducing the variance of the MLE $\hat{\beta}$. Therefore as many credit event companies as possible are desired. A *choice based* sampling method (also known as *case-control* or *endogenous stratified sampling*) uses all available, or some randomly selected observations for which a credit event took place and then selects randomly within the non-credit event companies. This yields a design that is consistent and efficient, but only with the appropriate correction. Two such choices of corrections are *Prior Correction* and *Weighting*. (King & Zeng, 2001)

Prior Correction

This is the simpler correction method of the two concerned with in this thesis, both conceptually and computationally. One initiates with the ordinary MLE calculation and once one has obtained the β_0 term, one corrects this factor by the ratio of the odds of a credit event within the sample and the true odds of a credit event for the whole population.

$$\hat{\beta}_{0,corr} = \hat{\beta}_0 - \log\left(\frac{1-\tau}{\tau} \cdot \frac{\bar{y}}{1-\bar{y}}\right), \quad (2.15)$$

where τ is the true fraction of credit events in the population and \bar{y} is the fraction of credit events within the sample. See Appendix B in King and Zeng (2001) for a proof of consistency of Prior Correction for logistic regression.

Weighting

Instead of maximising the usual log-likelihood function, a weighted log-likelihood function is introduced (King & Zeng, 2001),

$$\begin{aligned} \ln(L_w(\boldsymbol{\beta}|\mathbf{y})) &= w_1 \cdot \sum_{\{Y_i=1\}} \log(p_i) + w_0 \cdot \sum_{\{Y_i=0\}} \log(1-p_i) \\ &= - \sum_{i=1}^n w_i \cdot \log(1 + e^{(1-2y_i)\mathbf{x}_i\boldsymbol{\beta}}), \end{aligned} \quad (2.16)$$

where

^bThe regularity conditions include the following: the true parameter value β must be interior to the parameter space, the log-likelihood function must be thrice differentiable, and the third derivatives must be bounded. (Rodríguez, 2001)

$$\begin{aligned}
w_1 &= \tau/\bar{y} \\
w_0 &= (1 - \tau)/(1 - \bar{y}) \\
w_i &= w_1 Y_i + w_0(1 - Y_i).
\end{aligned}
\tag{2.17}$$

The expression of the weighted log-likelihood may look complex, but implementing the method is trivial because the weights w_i in Equation 2.17 can be calculated in advance. Any sophisticated logistic regression software will take weights as input. But, as discussed by Manski and Lerman (1977) and Xie and Manski (1989), the usual method for calculations of standard errors, based on the Information Matrix, is incorrect for Weighting. Furthermore King and Zeng (2001) mentions that the *"[...] problem is explained by the Information Matrix equality not holding under choice-based sampling."* and they proceed by illustrating the severity of the bias, and the increase of the bias in the number of left out non-event companies. They also present a solution, to apply the Huber-White (robust) standard error estimate. The Huber-White estimate of standard errors is the method utilised when calculating the Wald Statistics for Weighting in this thesis. See Freedman (2006) for an introduction.

The disadvantage of Prior Correction is that it is less robust than Weighting if the model is misspecified and with a large sample Weighting performs better. (Xie & Manski, 1989) However, when confident about the explanatory variables and the functional form of the model, Prior Correction is preferable. (King & Zeng, 2001) But it should be noted that Prior Correction is not always inferior to Weighting, as the latter is asymptotically less efficient. The illustration of this result, evident in a small sample situation, is attributed to (Scott & Wild, 1986; Amemiya & Vuong, 1987).

2.1.10 Two-Step Bias Correction

Illustration of Rare-Event Bias

First of all, there exists a bias inherent in any logistic regression estimation based on a sample not equal to the entire population; furthermore this bias is increasing in the rarity of the events. The bias in this thesis is towards underestimating the probability of experiencing credit events for companies, and thus overestimating the probability of surviving. To see this intuitively consider a model using one covariate with good discriminative power. If there are few credit events present in the credit event sample, only little information about the distribution of the covariate is obtained for these companies. Whereas there will be a more stable distribution for the non-credit event companies if there is a comparatively higher number of non-credit event observations. See Figure 2.1 for an illustration.

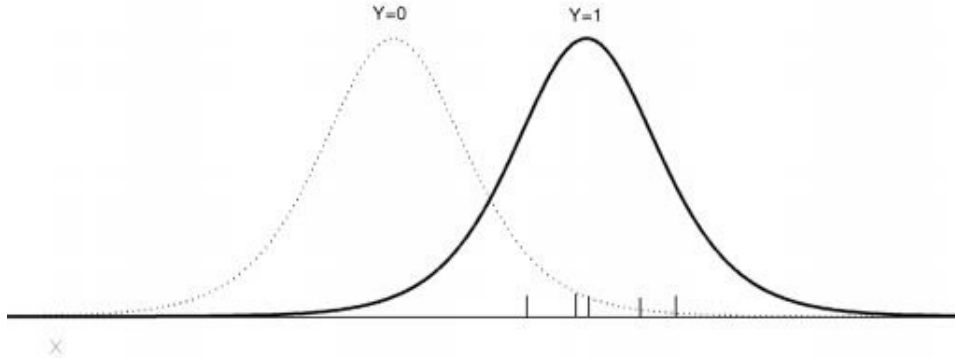


Figure 2.1: The few observed credit event companies ($Y = 1$) are marked as short vertical lines, along with the (solid) line for the density from which they were drawn. The many ($Y = 0$) non-credit event observations do not appear but their density is shown with the dotted line. (King & Zeng, 2001)

The left (non-credit event) distribution is dense and the right (credit event) distribution is relatively sparse, i.e. a rare event situation. Consider a classification between credit event companies and non-credit event companies based on the two given distributions. The maximum of the non-credit event population distribution will likely correspond well to the maximum of the non-credit event sample distribution; meanwhile, the minimum of the credit event sample is unlikely to correspond well to the minimum of the credit event population in a rare event situation. If the goal of the classification is to minimise the number of misclassifications, and type one and type two errors are equally important to consider, then the optimal value used for discrimination, i.e. the *cutoff*, will be very close to the minimum value within the credit event sample. As this minimum corresponds poorly to the minimum of the credit event population the model is biased towards classifying observations as non-credit events. This is the same as saying that any observation will be given too low probability of being classified as a credit event, and thus of course, also too high probability of being classified as a non-event company. This gives cause to the bias, and illustrates that the bias is increasing in the rarity of the events. Please see King and Zeng (2001) for a more thorough illustration of this effect.

Calculating and Compensating for the Bias in Beta

Following on McCullagh and Nelder (1989), who give an explicit estimation formula for the bias of any generalised linear regression model, King and Zeng (2001) gives proof for the special logistic regression case. Below the needed results for rare event bias correction of β are presented. Please see Section 15.2 and Appendix C of the former and latter named references for

the full derivation.

$$\text{bias}(\hat{\beta}) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\xi_i \quad (2.18)$$

where,

$$\xi = 0.5Q_{ii}[(1 + w_1)\hat{p}_i - w_1],$$

$$Q_{ii} \text{ are the diagonal elements of } Q = \mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}',$$

$$\mathbf{W} = \text{diag}\{\hat{p}_i(1 - \hat{p}_i)w_i\} \text{ and}$$

\hat{p}_i is the expected probability of an event based on the MLE estimates.

Note that this can be viewed as a simple weighted least square regression. The bias-corrected estimate is then $\tilde{\beta} = \hat{\beta} - \text{bias}(\hat{\beta})$ which by an approximation has variance $\text{Var}[\tilde{\beta}] = (n(n + k))2\text{Var}[\hat{\beta}]$ (King & Zeng, 2001). Where n is the number of observations and k is the number of columns in \mathbf{X} , i.e. the number of covariates in the model plus one. Be aware that the variance estimation is a crude approximation which works better for small values of β (McCullagh & Nelder, 1989).

Uncertainty in Beta

$\tilde{\beta}$ is preferable to $\hat{\beta}$ to use for calculation of the consistent expected probabilities, since $\tilde{\beta}$ is less biased and also, based on the estimation above, has lower variance than the MLE estimate $\hat{\beta}$. However, neither of these two estimates are optimal since they disregard the *uncertainty in beta* existing due to the fact that β is estimated rather than known. Furthermore, the uncertainty in beta is evident by a non-zero variance of the estimate. Since it is known that the MLEs of β in a logistic regression situation are asymptotically normally distributed, it is possible to mitigate the impact of the known uncertainty by utilising the law of total probability. (King & Zeng, 2001)

$$\mathbb{P}(Y_i = 1) = \mathbb{E}_{\tilde{\beta}}[\mathbb{P}(Y_i = 1|\tilde{\beta})] = \int \mathbb{P}(Y_i = 1|\beta^*)\mathbb{P}(\beta^*)d\beta^* \quad (2.19)$$

A simple way to calculate Equation 2.19 is to use a Monte Carlo scheme by drawing from the distribution of $\tilde{\beta}$.

2.1.11 Jackknife

The Jackknife method is commonly used to reduce bias and to calculate variances of cumbersome parameter estimates. It is especially useful if the parameter of interest has no explicit function. It is a re-sampling technique, part of a bigger family of methods known as *bootstrapping*, but the Jackknife method predates the more general bootstrapping methods. The method is

applied by simply calculating the *leave one out samples* of the vector of observed values.

Given a vector of observations \mathbf{x} containing values x_1, \dots, x_n , the Jackknife *leave one out samples* are the n vectors of length $n - 1$. The i 'th Jackknife sample vector is

$$\mathbf{x}[i] = \begin{cases} \{x_2, \dots, x_n\}, & \text{for } i = 1 \\ \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}, & \text{for } i = 2, \dots, n - 1 \\ \{x_1, \dots, x_{n-1}\}, & \text{for } i = n \end{cases} \quad (2.20)$$

Computing the variance of a parameter $\theta = g(\mathbf{x})$ is then easy. The n values $\theta[i] \equiv g(\mathbf{x}[i])$ are computed from which the estimate of the variance of θ follows as $\text{Var}(\hat{\theta}) = 1/(n - 1) \sum_i (\theta[i] - \theta_0)^2$, where θ_0 is the estimate of θ , retrieved by using the full data set. For large n , the jackknife estimate $\hat{\theta} = 1/n \sum_i \theta[i]$ is approximately normally distributed around the true parameter θ so it is possible to use the Jackknife method for constructing confidence intervals for θ . (Zhou, Obuchowski, & McClish, 2011)

2.1.12 Spearman's Rank Correlation

Spearman's rank correlation is a non-parametric statistic that can be used to test the strength of association between two variables. The statistic does not assume anything about the distribution of the variables except that the relationship between them is monotone and that the variables can be ranked ordinally. The statistic is simply defined and computed as the Pearson Correlation of the ranks in the data, i.e. the "*usual correlation formula*" computed on the ranks. The following formula can also be used to calculate Spearman's rank correlation in the case of distinct integer ranks,

$$\rho = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)} \quad (2.21)$$

where

ρ = Spearman's Rank Correlation,

d_i = The rank difference of paired observations i , for $i = 1, \dots, n$,

n = Number of observations in each data set.

Spearman's ρ is bounded by -1 and 1. Significance of the Spearman's rank correlation coefficient is based on a statistic which for large n is approximated by a normal or t-distribution. The details for the test are omitted but available in Kendall and Smith (1939). The p-value indicates the probability of seeing the observed correlation or stronger.

Cohen's standard, see Table 2.1, can be used to evaluate the Spearman's correlation coefficient to determine the strength of the relationship between the two variables. (Cohen, Cohen, West, & Aiken, 2002)

Spearman's ρ	Degree of Association
0.10-0.29	Small
0.30-0.49	Medium
0.50-	Large

Table 2.1: Cohen's Standard for Degree of Association

2.1.13 Winsorisation

Winsorisation is a method utilised to avoid and prevent the influence of outliers in data. The method has been used by previous researchers within the field of bankruptcy prediction, see Ohlson (1980) and Shumway (2001) for two examples. Before conducting a new experiment the 1 % and 99 % quantiles are computed for each covariate within the two samples of credit event companies and non-credit event companies. Values of covariates that fall outside of the corresponding sample quantiles are set to be equal to the value of the bounding quantile value. To the authors' knowledge, previous research has not performed winsorisation for credit event and non-credit event companies independently. There is a downside inherent in the proposed approach as it helps to separate the covariates which simplifies finding significant regression coefficients.^cOn the other hand, with a data set where the number of observations for the two groups differs greatly, there is an obvious downside to winsorising across the entire population. One introduces the risk of losing a lot of information for the smaller subset. Furthermore the negative impact of this effect is greatly enhanced by the discriminative efficiency of the covariate and the rarity of events in the population and in the sample. See Example 2.1.1 for an illustrative example.

Example 2.1.1. In preparation for the example 1000 i.i.d. normally distributed points are drawn with expected value 1 and standard deviation 0.5, these points are grouped in Group 1. 198 i.i.d. normally distributed points with expected value 3 and standard deviation 1 are also drawn which together with two outliers, manually put at 0, are grouped in Group 2. It should be noted that in a situation where the two groups are of equal size and the observations are identically distributed, the two winsorisation methods are expected to yield the same result. This would, however, translate to a situation in which a covariate would have very limited discriminative power, and where events would not be considered rare. In Figure 2.2 the top plot illustrate the distribution of the original example data. Note the 2 outliers at $x = 0$. The middle plot shows the same data but winsorised on covariate level, i.e. across both samples, at 2.5 % and 97.5 % levels. Note

^cAll winsorisations are naturally limited to only be used on the estimation set, as there is no feasible way to winsorise when the model is put into practice, and therefore, correcting the validation set would be a severe mistake.

now that the Group 2 data is greatly impacted by the large value of observations within Group 1. The winsorisation on the 97.5 % quantile impacts 15.0 % of all observations in Group 2. But the true outliers, located at 0 are not impacted. Meanwhile in the bottom plot the approach with group independent winsorisation yields a data set where the outliers are successfully altered and the data is not distorted asymmetrically between the two groups.

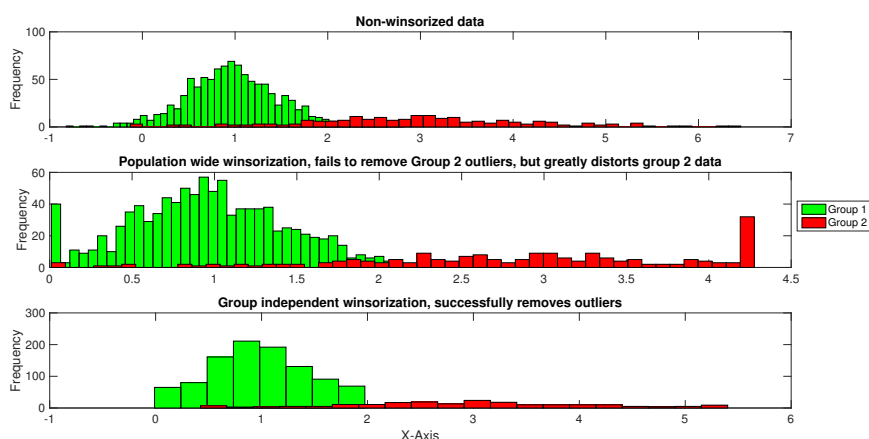


Figure 2.2: Winsorization Example

The group independent approach relies on the assumption that data is drawn from two distinct distributions, in this example it is obvious but in the data for this thesis such a conclusion is less trivial to draw.^d

2.2 Financial Theory

2.2.1 Corporate Bonds

Corporate bonds are securities that are sold by corporations in order to raise money today in exchange for promised future payments. The terms of the bond are described as part of the bond certificate, which indicates the dates and amounts of all future payments. The last payment is on the maturity date, the final repayment date. Bonds typically have two different types of payments, one is the interest payment of the bond, or coupon, and the other is the principal payment. The coupons are often paid periodically

^dAs indicated by both Welch's t-test and especially by Kolmogorov-Smirnov two sample test, the distributions of the covariates in this thesis are indeed, in many cases, distinct. Thus a group independent winsorisation not only avoids the problems illustrated above but is also the more theoretically sound method of the two.

with a pre-specified frequency (e.g. quarterly, semi-annual or annual) until the maturity date of the bond. The principal payment of the bond is repaid at the maturity date. (Berk & DeMarzo, 2013)

By investing in a bond the investor runs the risk of not being paid the promised future payments of the bond, as the corporation may not pay back the full amount. The risk of default of a bond is known as the bond's credit risk. (Berk & DeMarzo, 2013)

2.2.2 Credit Default Swaps

In a credit default swap (CDS), the buyer pays a periodic premium to the seller of the swap and receives a payment from the seller if the underlying security (often bond) defaults. The contracts allow market practitioners to transfer the credit risk of a company. Traditionally, CDS spreads represent the fair insurance price for the credit risk of a company. CDS contracts are written between counterparties and traded over-the-counter (OTC). A buyer or seller who wants to unwind a position can't sell or buy the contract on an exchange like stocks, but is instead forced to enter into an offsetting CDS contract with a possibly new counterparty. (Berk & DeMarzo, 2013, p. 728-729). The contractual nature of CDS contracts makes them less influenced by convenience or liquidity factors than bond assets. (Arakelyan & Serrano, 2012)

2.2.3 Financial Statements

Financial statements are accounting reports periodically (usually quarterly and annually) issued by corporations. They present a snapshot and summarise past and current information of a corporation's financial status. Public companies (i.e. companies traded on a stock exchange) are forced to submit an annual report with their financial statements to their shareholders each year. Private companies (i.e. companies not traded on any public exchange) often prepare and publish the same type of reports, even though they are not obliged to. (Berk & DeMarzo, 2013, p. 22)

Public corporations have to present four financial statements; the balance sheet, the income statement, the statement of cash flows, and the statement of changes in shareholders' equity. The authors assume that the reader has basic knowledge of financial statements and refer to Berk and DeMarzo (2013) for a thorough discussion of the subject.

2.2.4 Financial Ratios

Financial ratios play an important role in financial reporting. A financial ratio consists of a numerator and denominator relating two financial amounts. The financial amounts can be from any of the four financial statements that the corporation issues. Financial ratios aid in the benchmarking process

of a corporation's performance as they, by introducing comparability, help to identify problem areas within a corporation's operations, liquidity, debt position, profitability, etc. (Faello, 2015)

Financial ratios not only provide information of past performance of a company but many also interpret them as guidance of where a firm is heading. For example, negative trends, or states, of financial ratios could indicate that a firm is in decline and provide insights into the prediction of corporate failure. (Faello, 2015)

2.2.5 Credit Rating Agencies and Credit Ratings

A central issue in finance is the lender's uncertainty concerning whether a borrower will fulfill all the contractual obligations of a loan or not. This can be thought of in terms of asymmetric information, i.e. the borrower knows more of its capabilities and financial status than the lender does. Consequently, the lender will, prior to extending a loan, want to gather information about prospective borrowers, in order to determine their creditworthiness. Following the extension of a loan, the lender will want to monitor the borrower's actions, and creditworthiness, to be reassured that the contractually obliged repayments are not in jeopardy.

Credit rating agencies provide a means to reduce the named asymmetric information inherent in financial markets. The "Big Three" agencies are Standard & Poor's (S&P), Moody's and Fitch Group. After collecting information about the bond issuers, the credit rating agencies offer judgements, called "opinions"^e, about the creditworthiness of bonds, corporations, and sovereigns. The judgements are in the form of ratings of which Standard & Poor's are the most well-known and have the structure of AAA, AA, A, BBB, BB, ..., C, D (including +/-). (White, 2010)

2.2.6 Definition of Credit Event

This thesis uses Moody's definition of default which is applicable to debt or debt-like obligations (e.g. bonds, swap agreements, etc.) (Moody's Investor Services, 2016). Moody's has four events that fall under their definition of default,

- a missed or delayed disbursement of a contractually-obliged interest or principal payment (excluding missed payments cured within a con-

^e"The rating agencies prefer that word because it allows them to portray themselves as publishers, akin to the publishers of newspapers, and thereby gain the protection of the First Amendment of the U.S. Constitution when they are sued by unhappy investors (e.g., who claim that they were injured by ratings that were subsequently shown to be overly optimistic) or by issuers (e.g., who claim that they were injured by overly pessimistic ratings)" (White, 2010)

tractually allowed grace period), as defined in credit agreements and indentures;

- a bankruptcy filing or legal receivership by the debt issuer or obligor that will likely cause a miss or delay in future contractually-obliged debt service payments;
- a distressed exchange whereby 1) an obligor offers creditors a new or restructured debt, or a new package of securities, cash or assets that amount to a diminished financial obligation and 2) the exchange has the effect of allowing the obligor to avoid a bankruptcy or payment default in the near future;
- a change in the payment terms of a credit agreement or indenture imposed by the sovereign that results in a diminished financial obligation, such as a forced currency re-denomination (imposed by the debtor, himself, or his sovereigns) or a forced change in some other aspect of the original promise, such as indexation or maturity.

2.3 Earlier Models

2.3.1 Altman Z-Score

Professor Edward I. Altman introduced the first multivariate bankruptcy prediction model in 1968 (Altman, 1968). His model, which is now more known as the Z-score model, was a breakthrough in the academic field of bankruptcy prediction, based on financial ratios and other variables to systematically assess credit qualities. In his 1968 paper, Altman uses a data set containing 66 American manufacturing companies. In his data set, half of the companies had filed for bankruptcy during the period 1946-1965. The non-bankrupt companies were chosen in, what Altman describes as, a *stratified random basis* (comparable to *case-control* as described in Section 2.1.9), based on industry, asset size and that they were still existent in 1966. The financial ratios needed are obtained from financial statements one reporting period prior to bankruptcy. Altman uses a total of 22 ratios, some from past studies and others introduced by him as likely successful predictors of financial distress. From the original 22 ratios, five are included in his model. In order to arrive at the final ratios, Altman's procedure combines the following four points, "(1) observation of the statistical significance of various alternative functions, including determination of the relative contributions of each independent variable; (2) evaluation of inter-correlations among the relevant variables; (3) observation of the predictive accuracy of the various profiles; and (4) judgment of the analyst." (Altman, 1968).

Altman utilises Multiple Discriminant Analysis (MDA) to obtain his final model. MDA is a statistical technique which classifies observed response

variables into predefined groups dependent on the characteristics of the individual observation. When using the resulting MDA model a company specific score can be calculated by use of a discriminant function. Comparing the resulting score to calibrated or predefined values yields a classification. The procedure in his 1968 paper results in the discriminant function in Equation 2.22,

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5 \quad (2.22)$$

where,

$$X_1 = WC/TA = \text{Working Capital/Total Assets}$$

$$X_2 = RE/TA = \text{Retained Earnings/Total Assets}$$

$$X_3 = EBIT/TA = \text{Earnings Before Interest \& Taxes/Total Assets}$$

$$X_4 = MCAP/TL = \text{Market Value Equity/Book Value of Total Debt}$$

$$X_5 = Rev/TA = \text{Sales/Total Assets}$$

Furthermore, variable X_1 to X_4 should be inserted as percentage values (i.e. 1 % is written as 1 rather than 0.01) and X_5 is inserted in the normal decimal way (i.e. 1 % is written as 0.01 rather than 1). Due to this obvious practical confusion a more convenient version of the model has emerged and is presented in Equation 2.23, where all values are used in the normal decimal way (i.e. 1 % is written as 0.01 rather than 1). This version is also suggested by Altman (2000). It is important to note that Altman's Total Assets variable uses Tangible Assets if available.

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5 \quad (2.23)$$

Altman introduces a *cutoff* to enable use of the model for discrimination between high and low risk of bankruptcy, where a Z-score higher than the cutoff indicates a high risk and a Z-score lower than the cutoff is associated with a low risk. Using his Z-score and a cutoff of 2.67, Altman achieves 95 % correct classification of his 66 in-sample companies. Furthermore, he manages to predict 24 of 25 bankrupt companies correctly in an out-of-sample test. In addition, he applies the model to 66 distressed, but still not bankrupt companies, and among these he manages to predict 79 % companies correctly, i.e. as non-bankrupt.

Altman appreciates that all companies are unlikely to be easily divided into two mutually exclusive groups based on a cutoff. In an attempt to mitigate the risk of misclassification he expands his model to include a *grey zone*, i.e. a non-certain zone. Therefore, after empirically testing the model, Altman suggests that a company with a Z-score below 1.81 should be considered bankrupt and companies with a Z-score above 2.99 should be considered

“non-bankrupt”. A company with a Z-score between 1.81 and 2.99 should be put in the *grey zone*, or *zone of ignorance*. (Altman, 1968)

The Altman Z-score model has since 1968 been revisited and several newer versions have been constructed, such as the Z' and Z" models which address US Private Manufacturing and US Non-Manufacturing and Foreign Firms respectively. (Altman, 2000) These newer models are not considered in this report.

2.3.2 Ohlson O-Score

Professor James Ohlson developed a bankruptcy prediction model based on logistic regression in 1980. Ohlson chose to use logit analysis as he wanted to avoid well-known problems associated with MDA. In his 1980 report he lists three of these problems, (1) The covariance matrix of the predictors should be the same for failed and non-failed firms, and a requirement of normally distributed predictors rules out the applicability of dummy variables, (2) The score from the MDA model has no intuitive interpretation since it is in essence an ordinal ranking tool, (3) The matching procedure between failed and non-failed firms, typically utilised for MDA, is often based on size, industry, and other measures which Ohlson argues tend to be arbitrary and could instead be considered as predictors. (Ohlson, 1980)

Ohlson uses a sample of 105 bankrupt and 2,058 non-bankrupt industrial firms from 1970 to 1976. In his paper he estimates three models based on nine independent variables. The first model predicts bankruptcy within one year, the second within two years, and the third within one or two years. His first model is of interest in this thesis and it is presented in Equation 2.24. The O-score is then converted to a probability with Equation 2.25.

$$O = -1.32 - 0.407X_1 + 6.03X_2 - 1.43X_3 + 0.0757X_4 - 2.37X_5 - 1.83X_6 + 0.285X_7 - 1.72X_8 - 0.521X_9 \quad (2.24)$$

$$p = \frac{1}{1 + e^{-O}} \quad (2.25)$$

Where,

$X_1 = Size = \log(\text{Total Assets}/\text{GNP price-level index})$, where 1968 is used as a base value of 100 for the index

$X_2 = TL/TA = \text{Total Liabilities}/\text{Total Assets}$

$X_3 = WC/TA = \text{Working Capital}/\text{Total Assets}$

$X_4 = CL/CA = \text{Current Liabilities}/\text{Current Assets}$

$X_5 = NI/TA = \text{Net Income}/\text{Total Assets}$

$X_6 = FU/TL = \text{Funds Provided by Operations}/\text{Total Liabilities}^f$

$X_7 = INTWO =$ One if NI was Negative for the Last Two Years,
Zero Otherwise

$X_8 = OENEG =$ One if TL > TA, Zero Otherwise

$$X_9 = CHIN = \frac{NI(t) - NI(t-1)}{|NI(t)| + |NI(t-1)|}$$

cutoff = 0.50

WC/TA, CL/CA and INTWO are not significant in Ohlson's paper, but he still included them. Ohlson had 96.12 % correct predictions for his data sample when he used 0.5 as the cutoff probability. Ohlson did not perform any out-of-sample testing. (Ohlson, 1980)

2.4 Model Evaluation Theory

2.4.1 Classification Table

One common way to summarise the predictive power of a logistic regression model is by use of a *classification table* (or *confusion matrix*). The table orders the actual binary outcome ($y = 0$ or 1) together with the prediction given by the model ($\hat{y} = 1$ or 0) in a 2×2 matrix. The prediction of $\hat{y} = 1$ is given if the $p_i > \text{cutoff}$ and $\hat{y} = 0$ if $p_i \leq \text{cutoff}$. If the model predicts a company to survive ($\hat{y} = 0$) but it actually has failed, then it is called a False Negative. If a company survived ($y = 0$) while the prediction is positive ($\hat{y} = 1$), then it is called a False Positive. A company that survived ($y = 0$) that has been correctly classified is called a True Negative result and a correct prediction of a failing company is called a True Positive. By changing the cutoff probability, cutoff, the predictions would change and therefore also the classification table would change. The choice of cutoff probability should be such that the overall cost of misclassification is minimised. (Agresti, 2007, p.142-143)

Two crude measures of model performance are the Positive and Negative Predictive Values, defined as the empirical estimates of the probabilities $P(y = 1|\hat{y} = 1)$ and $P(y = 0|\hat{y} = 0)$. Two other useful measures of predictive power are sensitivity $P(\hat{y} = 1|y = 1)$ and specificity $P(\hat{y} = 0|y = 0)$. The overall proportion of correct classification can also be used as a summary of predictive power. It is defined as: $P(\text{correct classification}) = P(y = 1 \cap \hat{y} = 1) + P(y = 0 \cap \hat{y} = 0)$, and can be thought of as a weighted average of sensitivity and specificity.

^fIn this thesis Cash From Operations will be used as a proxy for Funds Provided by Operations

		Actual Outcome		
		$y = 1$	$y = 0$	
Model Prediction	$\hat{y} = 1$	True Positive (TP)	False Positive (FP)	Pos. Pred. Value
	$\hat{y} = 0$	False Negative (FN)	True Negative (TN)	Neg. Pred. Value
		Sensitivity	Specificity	

Table 2.2: Classification Table; Where Sensitivity = $\frac{TP}{TP+FN}$, Specificity = $\frac{TN}{FP+TN}$, Positive Predictive Value = $\frac{TP}{TP+FP}$ and Negative Predictive Value = $\frac{TN}{FN+TN}$

2.4.2 Cumulative Accuracy Profile

In order to assess the discriminative power of a model the method of Cumulative Accuracy Profile (CAP) can be used as a visual tool. To construct the CAP curve, companies are ranked in increasing order of credit quality according to their score from the model. “The CAP curve is constructed by plotting the fraction of all defaults that occurred among borrowers rated x or worse against the fraction of all borrowers that are rated x or worse” (Löffler & Posch, 2007, p.148-151). A default prediction model that performs well should assign the highest probabilities of defaults in the sample to the companies that have defaulted. For a model that has no discriminative power, i.e. a model no better than guessing, the CAP curve is expected to form a “45 degree line”, a *random assignment line*.

Example 2.4.1. In Figure 2.3 an example of three CAP curves is presented. In the example there are 20 companies that have experienced a credit event (credit event companies) and 80 companies that have not experienced credit events (non-credit event companies). The *perfect model* assigns the 20 highest probabilities to the credit event companies. The *acceptable model* manages to find all 20 credit event companies after going through the 40 worst companies according to the model’s ranking. A model that has no discriminative power is expected to have slope 1, *45 degree line*.

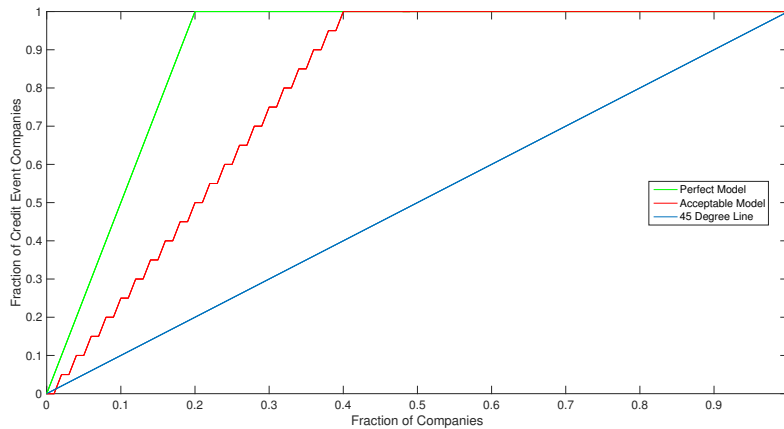


Figure 2.3: Example of Cumulative Accuracy Profiles

2.4.3 Area Under Curve (or Accuracy Ratio)

The information provided by the CAP curve can in large be captured by a single number, the Area Under Curve (AUC). The AUC for a model is defined as the ratio between two areas, (1) The area between the CAP-curve of the model and the *random assignment line*, and (2) The area between the *perfect model's* CAP-curve and the *random assignment line*. The *acceptable model* in Example 2.3 has an $AUC = 0.75$. The AUC is always bounded by $[-1, 1]$. The AUC for a model should be above zero as it otherwise is outperformed by a model that randomly assigns ranks. Löffler and Posch (2007) state that credit rating systems that are used in practice have a typical AUC between 0.5 and 0.9. AUC and CAP should be used carefully as they do not discriminate between the cost of type I and type II errors.

2.4.4 Receiver Operating Characteristic

Sensitivity and specificity and other measures of classification performance computed from the classification tables depend on a single cutoff probability. A better and more complete description of classification accuracy of a model is the area under the Receiver Operating Characteristic curve (ROC-curve). The ROC-curve plots the sensitivity and (1 - specificity) for a range of cutoff probabilities. This method has according to Hosmer, Lemeshow, and Sturdivant (2013) become the standard for evaluating a fitted model's discriminative ability. Understanding the construction of the ROC-curve yields an intuitive interpretation of the choice of cutoff probability, for discriminative models, as the intersection of the sensitivity and (1 - specificity) curves. This cutoff is furthermore a common choice in practice. (Hosmer et al., 2013)

The area under the ROC-curve, known as ROC, ranges from 0.5 to 1.0, where 0.5 indicates no discrimination and 1.0 indicates perfect discrimination. According to Hosmer et al. (2013) there is no area under the ROC-curve that indicates a clear difference between a good or bad model but a suggestion from Hosmer et al. (2013) is to use the rule of thumb in Equation 2.26.

$$\text{If} = \begin{cases} ROC = 0.5 & \text{No discrimination.} \\ 0.5 < ROC < 0.7 & \text{Poor discrimination.} \\ 0.7 \leq ROC < 0.8 & \text{Acceptable discrimination.} \\ 0.8 \leq ROC < 0.9 & \text{Excellent discrimination.} \\ ROC \geq 0.9 & \text{Outstanding discrimination.} \end{cases} \quad (2.26)$$

The ROC-curve is a tool similar to the CAP, both show sensitivity on the y-axis but against different x-axes. The similarity between the ROC-curve and CAP is further reflected in that there is a linear relationship between the area under the CAP-curve, AUC, and the area under the ROC-curve, ROC,

$$AUC = 2 \cdot ROC - 1 \quad (2.27)$$

Please see Löffler and Posch (2007, p.151-152) for more information on the subject.

3

Data

Data! Data! Data! I can't make bricks without clay!

– Sir Arthur Conan Doyle

All models considered in this report, need financial data that are made available through financial statements for credit event and non-credit event companies. The choice of which annual statement to use for credit event companies is straightforward, the one from one year prior to the credit event feels the most natural. The decision concerning which annual statement to include for the non-credit event companies is more ambiguous. The goal of this data section is to resolve the ambiguity and to construct a credit event sample and a non-credit event sample so that the characteristics of credit events are able to be captured, rather than sample differences. The two samples are, within reason, tried to be made as similar as possible in terms of asset size, industry classification and year-distribution of the annual statements.

3.1 Credit Event Sample

Through Moody's annual "Corporate Default and Recovery Rates" reports (for an example see Services (2015)), access is granted to credit events that have occurred in the United States between the years 2002-2014. For credit events prior to 2002, accounting data on the trading platform Bloomberg is too sparse to prove useful. The data gathering from Moody's reports results in 736 credit events. Some events are registered for the same companies, the later occurring events are disregarded due to the risk of temporal dependence this can introduce between the observations.^aThis filtration results in a sample of companies containing 654 credit events.

^aIndependence among observations is for example needed for calculating the MLEs of coefficients and in extension of the probabilities. Furthermore, a company that has experienced a credit event previously may act differently compared to one which has not.

From the Moody's reports the names of the companies and the years of the credit events are obtained. For some reports the industry of the company, the initial default event type and the month of the default are also listed. The Bloomberg terminal is then used to gather accounting data for the companies. Unfortunately the names of the companies in Moody's reports do not exactly or uniquely match the names of companies in the Bloomberg terminal. Preferable would be if the companies in the reports also have BBG-tickers available, but they do not. This forces a manual task where each credit event is checked for the unique matching company with the correct financial statement. Companies are used if and only if,

- a unique company can be found;
- accounting data is available for the year prior to the year of the credit event;
- the company is non-financial and not a real estate company or a real estate investment trust (REIT);
- the company is not a subsidiary of, and has not been acquired by, another company in the sample;
- the company has not been charged for fraudulent accounting practice in the time period investigated.

345 companies out of the 654 credit events in the sample are removed using the five criteria above.^bThe final sample of credit events thus therefore consists of 309 companies.

The criteria that Bloomberg must have data on the companies is of course unfortunate as this introduces a sort of undeniable bias. The bias, however, can be viewed from multiple standpoints. The two most important implications, for the purposes of this report, are: (1) A bias towards inclusion of larger and more popular companies and (2) A bias towards exclusion of earlier defaulted companies. Both of these biases can however be seen as positive. As the potential bias is towards more important companies from the practitioners' points of view. As a final note, the criteria is necessary, since the data gathering process is limited to use of the Bloomberg database.

Therefore, the model can only be calibrated to predict the first occurrence of a credit event for a company. An obvious limitation in the data gathering is that it was not tractable to check for credit events prior to 2002, but since such observations are unused in the calibration, no temporal dependence issues are introduced by the credit event filtration.

^bBloomberg classified 43 companies as Financials or REITs and were thus removed. 260 companies for which we couldn't find a unique company and/or accounting data were also removed. The remaining 42 companies are removed for other reasons which include, but are not limited to, (1) Being subsidiaries of or have been acquired by other companies that have registered credit events at an earlier time or (2) The company has been caught for fraudulent accounting practices.

3.2 Non-Credit Event Sample

The non-credit event sample should ideally be all those companies overlooked by Moody's but that have not yet experienced any credit event. Such a non-credit event sample is unfortunately not obtainable and a proxy is therefore constructed. The full non-credit event sample consists of 1,002 companies. For full disclosure, see Appendix A which concerns the construction of the proxy, i.e. the non-credit event sample.

3.2.1 Selection Bias

All of the companies in the credit event sample are retrieved from Moody's database, thus all have been, or are currently, tracked by Moody's. Meanwhile, the non-credit event companies are chosen so that an issuer rating from either Moody's, S&P or Fitch exist in the Bloomberg Database. This introduces a selection bias as the non-credit event companies could potentially not share the characteristics of "being tracked by Moody's". By including S&P- and Fitch-data for non-credit event companies the asset sizes of the two cohorts are more similar in terms of asset size. This naturally increases the number of observations, which is considered beneficial.

3.3 Final Adjustments to the Samples

3.3.1 Sample Differences in Total Assets

After observing the asset size data for the two cohorts it appears that the credit event sample has 30 companies whose total asset size is above \$5 bn. The non-credit event data set has 392 companies above the same threshold. To include all companies from the two samples unabashedly would be a terrible mistake, because the two total asset size distributions for the samples would greatly differ. To avoid this complication, a restriction is put on the total asset sizes for the credit event companies at \$10 bn. It is furthermore noted that above \$5 bn the credit event data is worrisome sparse, with only 11 companies between \$5 bn and \$10 bn. Therefore, 11 companies with total assets size between \$5bn and \$10bn are randomly drawn from the non-credit event sample. The asset size distribution matching above is made as large companies can behave quite differently in times of crisis (Vassalou & Yuhang, 2004), which can for example be explained through disposals of subsidiaries.

After the sample differences in total assets are adjusted for, the final data set consists of 292 credit event companies and 619 non-credit event companies.

3.3.2 Synchronizing the Year Distributions

As mentioned in the introduction to the Data section, the non-credit event sample should share characteristics with the credit event sample. More specifically, the years from which annual statements are taken within the non-credit event sample are desired to be unbiased with respect to the distributions of years and industries within the credit event sample.

This is obtained by modelling the distribution of years for each industry within the credit event sample as independent multinomial distributions. I.e. for an arbitrary credit event company the industry specific distribution, of when the credit event occurred, is assumed to be multinomial with one category for each of the years 2002-2014. The MLE of the probabilities of experiencing a credit event in any of the possible years is simply the ratio of the number of credit-events occurring in that year, and industry, divided by the total number of credit events for the same industry. For an arbitrary non-credit event company, a draw is made from the recently constructed industry specific year distribution. This yields approximately the same distribution of years for annual statements among non-credit event companies and credit event companies. A key assumption is that non-credit event and credit event companies are from the same industry specific populations. By extension approximately equal sample wide year distributions are obtained by aggregation.

In Figure 3.1 the industry specific year distributions are illustrated. In Figure 3.2 the aggregated year distribution and sector distribution of the whole sample is illustrated.

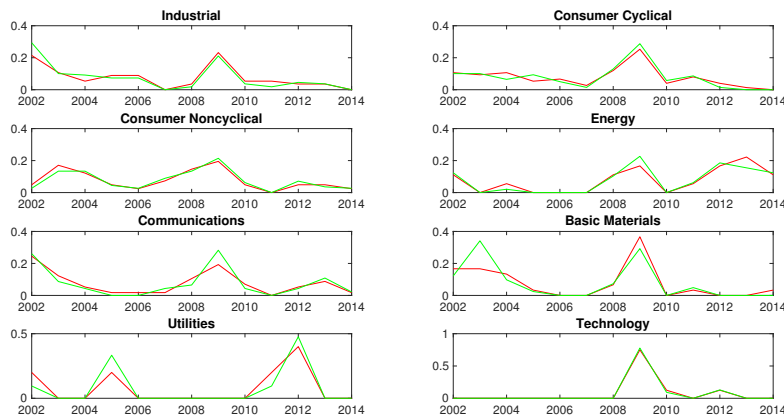


Figure 3.1: Year distribution for all of the sectors in the sample

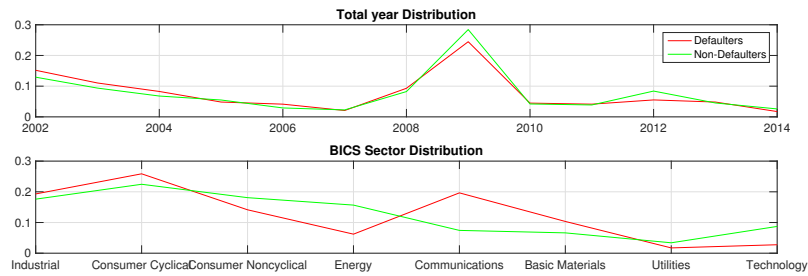


Figure 3.2: Aggregated year distribution and BICS sector distribution

3.4 Ratios

The ratios^c that will be considered in the model building phase are listed in Table 3.1. The ratios are taken from Altman's and Ohlson's models together with an aggregation of ratios that are mainly obtained from Beaver (1966) and market practitioners^d. Since logistic regression assumes linearity of covariates in the output (log-odds), the logarithm is taken of ratios that have strictly positive support. The Bloomberg formulae that are used to extract the financial data are available in Appendix B along with definitions of all ratio constituents.

^cIn all ratios where Total Assets are used, it is in first hand attempted to use Tangible Assets, if such a data point exist. The same procedure is also followed by Altman.

^dSpecial thanks to Ingvar and Pia at Swedbank.

Type	#	Ratio	Description	Abbreviation	From
Cash Flow	1	Cash from Operations to Total Liabilities	Cash From Operations/Total Liabilities	CFO/TL	Ohlson 1980
	2	Cash Ratio 1	Cash and Near Cash/Current Liabilities	Cash Ratio 1	Beaver 1966
	3	Cash Ratio 2	Cash, Cash Eq. & STI /Current Liabilities	Cash Ratio 2	-
	4	Cash to Sales	Cash and Near Cash/Revenue	CASH/Rev	Beaver 1966
	5	Cash to Total Assets	Cash and Near Cash/Total Assets	CASH/TA	Beaver 1966
	6	Free Cash Flow to Total Liabilities	Free Cash Flow/Total Liabilities	FCF/TL	-
	7	Interest Service Cover Ratio	Free Cash Flow/Financial Expenditure	IntSerCR	-
Profitability	8	EBIT margin	EBIT/Revenue	EBIT margin	-
	9	EBIT to Total Assets	EBIT/Total Assets	EBIT/TA	Altman 1968
	10	EBIT to Total Interest Expense	EBIT/Total Interest Expense	EBIT/TIntExp	-
	11	EBITDA margin	EBITDA/Revenue	EBITDA margin	-
	12	EBITDA to Net Debt	EBITDA/Net Debt	EBITDA/ND	-
	13	EBITDA to Total Debt	EBITDA/Total Debt	EBITDA/TD	-
	14	EBITDA to Total Interest Expense	EBITDA/Total Interest Expense	EBITDA/TIntExp	-
	15	Net Income margin	Net Income/Revenue	NI/Revenue	Beaver 1966
	16	Net Income to Total Assets	Net Income/Total Assets	NI/TA	Ohlson & Beaver
	17	Net Income to Total Debt	Net Income/Total Debt	NI/TD	Beaver 1966
	18	Net Income to Total Equity	Net Income/Total Equity	NI/TE	-
	Debt, Liability & Equity	19	Net Income to Total Liability	Net Income/Total Liability	NI/TL
20		Retained Earnings to Total Assets	Retained Earnings/Total Assets	RE/TA	Altman 1968
21		Intangibles to Total Equity	Intangibles/Total Equity	INT/TE	-
22		Long Term Debt to Total Assets	Long Term Debt/Total Assets	LTD/TA	-
23		Long Term Debt to Total Debt	Long Term Debt/Total Debt	LTD/TD	-
24		Long Term Debt to Total Invested Capital	Long Term Debt/Total Invested Capital	LTD/TotInvCap	-
25		OENEG	1 if TL > TA, 0 otherwise	OENEG	Ohlson 1980
26		Short Term Debt to Total Debt	Short Term Debt/Total Debt	STD/TD	-
27		Short Term Debt to Total Invested Capital	Short Term Debt/Total Invested Capital	STD/TotInvCap	-
28		Solvency	Total Equity/Total Assets	Solvency	-
29		Solvency Without Goodwill	(TE - Goodwill)/(Total Assets - Goodwill)	SwG	-
Size		30	Total Equity to Long Term Debt	Total Equity/Long Term Debt	TE/LTD
	31	Total Equity to Net Debt	Total Equity/Net Debt	TE/ND	-
	32	Total Equity to Short Term Debt	Total Equity/Short Term Debt	TE/STD	-
	33	Total Equity to Total Debt	Total Equity/Total Debt	TE/TD	-
	34	Total Equity to Total Liabilities	Total Equity/Total Liabilities	TE/TL	-
	35	log(Total Liabilities to Total Assets)	log(Total Liabilities/Total Assets)	TL/TA	Ohlson & Beaver
	36	Size	log(Total Assets/GDP Price Index)	Size	Ohlson 1980
	37	log(Current Assets to Revenue)	log(Current Assets/Revenue)	CA/Revenue	Beaver 1966
	38	log(Current Assets to Total Assets)	log(Current Assets/Total Assets)	CA/TA	Beaver 1966
	39	log(Current Liabilities to Current Assets)	log(Current Liabilities/Current Assets)	CL/CA	Ohlson & Beaver
	40	log(Current Liabilities to Total Assets)	log(Current Liabilities/Total Assets)	CL/TA	Beaver 1966
	41	log(Quick Ratio)	log(Quick Ratio)	QR	-
Liqud Assets	42	Working Capital to Revenue	Working Capital/Revenue	WC/Revenue	Beaver 1966
	43	Working Capital to Total Assets	Working Capital/Total Assets	WC/TA	Altman, Ohlson & Beaver
	44	Accounts Payable Turnover	Accounts Payable Turnover	APT	-
	45	Accounts Receivable to Accounts Payable	Accounts Receivable/Accounts Payable	AR/AP	-
	46	Accounts Receivable to Revenue	Accounts Receivable/Revenue	AR/Revenue	Beaver 1966
	47	log(Accounts Receivable Turnover)	log(Accounts Receivable Turnover)	ACT	-
	48	Cash Conversion Cycle	Acc. Rec. T. + Inv. T. - Acc. Pay. T	Cash C. C.	-
	49	log(Inventory Turnover)	log(Inventory Turnover)	IT	-
	50	Inventory Turnover to Working Capital	Inventory Turnover/Working Capital	IT/WC	-
	51	log(Quick Assets to Current Liabilities)	log((Acc. Rec. + Cash and Near Cash)/CL)	QA/CL	Beaver 1966
	52	log(Quick Assets to Sales)	log((Acc. Rec. + Cash and Near Cash)/Revenue)	QA/Revenue	Beaver 1966
	Activity	53	log(Quick Assets To Total Assets)	log((Acc. Rec. + Cash and Near Cash)/TA)	QA/TA
54		Revenue to Total Assets	Revenue/Total Assets	Revenue/TA	Altman & Beaver
55		log(Revenue to Total Debt)	log(Revenue/Total Debt)	Revenue/TD	-
56		log(Market Capitalization to Total Liabilities)	log(Market Capitalization/Total Liabilities)	MCAP/TL	Altman 1968
57		Net Income to Market Value Total Assets	Net Income/Market Value Total Assets	NI/MTA	-
58		Total Liabilities to Market Value Total Assets	Total Liabilities/Market Value Total Assets	TL/MTA	-
59		Working Capital to Market Capitalization	Working Capital/Market Capitalization	WC/MCAP	-
60		Change In Net Income (CHIN)	(NI(t) + NI(t-1))/(abs(NI(t))+abs(NI(t-1)))	CHIN	Ohlson 1980
61		Current Asset Quality to Current Liability Quality	Accts Rec/Current Assets*Current Liabilities/Accts Pay	CAQ/CLQ	-
62		INTWO	1 if NI negative past 2 years, 0 otherwise	INTWO	Ohlson 1980
63		Net Sales Change	(Rev(t) - Rev(t-1))/Rev(t-1)	Net Sales Change	-

Table 3.1: Illustration of ratios that are used

3.5 Estimation & Validation Sets

The final data set of credit and non-credit event companies is by randomization separated into two sets of equal size, namely the estimation- and validation set. The estimation set is used for building and calibrating models and the validation set is only used to evaluate the out-of-sample performance of the different models.

3.6 Market Data

As a final part of this thesis, two types of rank consistency for the final model's output are examined. Firstly, rank consistency to credit ratings provided by S&P. Secondly, rank consistency to market observed CDS spreads.

Both types are checked using data from the US and the Nordic market.

3.6.1 Rating Data

Similarly to the non-default data set the EQS tool in Bloomberg is used to retrieve a population of rated companies. The filters that are applied for the US and Nordic markets are listed in Appendix C. For the US market 1,054 companies are found and the same number for the Nordic market is 35.

3.6.2 CDS Data

US Companies

The source of the data is again the Bloomberg Terminal, but now the GCDS tool is utilised, see Appendix C for the details. For the US, 436 CDS contracts are found. For each contract the corresponding reference entity's latest annual statement is selected, from which the financial ratios are extracted. For all of the CDS contracts the spreads are taken from the reference entities financial statements' announcement dates and from dates in the following 4 weeks.[°]In order for the company to be selected the latest annual statement has to be from 2014 or 2015. Furthermore, if the company is missing 2 or more of the 5 CDS spreads then the company is removed from the sample. 115 companies remains after applying these restrictions. The mean of these obtained spreads is then computed for each company, which the ranking is based on.

Nordic Companies

For the Nordic region 29 CDS contracts are found. 23 of these are selected as they all have annual statement issued for 2015. The gathering process of the CDS spreads is unfortunately not possible to automate, as for the US companies. Consequently, the simplified approach for each of the 23 CDS contracts is instead to find the four end of week spreads (the last traded spread of each week) following each company's annual announcement date. As for the US CDS data a ranking is then constructed after computing the mean of the CDS spreads for each company.

[°]A company that has an annual statement announcement date at the 1st of February the CDS spreads are attempted to be retrieved for the 1st, 8th, 15th, 22nd of February and 1st of March (That is the announcement date +0, +7, +14, +21 and +28).

4

Method

Data do not give up their secrets easily. They must be tortured to confess.

– Jeff Hopper

This section concerns, (1) How a model is built from the data described in the Data section; (2) How different models will be evaluated against each other; and (3) How model output is tested for rank consistency with CDS and Rating data.

4.1 Model Building

In order for a model to perform well, one needs ratios that can discriminate between credit event and non-credit event companies. A logistic regression approach is chosen, but the choice of covariates within the model building phase is far from trivial. The goal of any model building method is to find the “best” possible model based on the available resources. In order to achieve this goal one must have a plan for selection of explanatory variables as well as a sound method for assessing the performance of the model. As Hosmer et al. (2013) put it: “Successful modeling of a complex data set is part science, part statistical methods, and part experience and common sense.”.

4.1.1 Flexible Data Set

All companies in the estimation and validation sets have annual statements available, but that does not imply that all ratios from Table 3.1 are available for analysis. For every additional ratio considered for analysis, all companies that lack the additional ratio need to be temporarily excluded. If a limitation is put to include only the companies that have all 63 ratios available, then the number of credit event companies available for analysis, would be reduced from 292 to 66. This is of course unwanted as the final model is not likely to

contain all ratios. By the use of a *flexible data set* it is possible to retain as many companies, and therefore, as much information as possible, for each step in the analysis. The implemented *flexible data set* approach is more dynamic and theoretically sound but in practice slightly more demanding. In theory 2^{63} subsets of the final credit event and non-credit event populations are considered, as each company either has the ratios of interest or they do not.

The *flexible data set* introduces no bias in the calibration or testing phases, but one could argue that there is a bias introduced in the model building phase, this is discussed briefly below. For a pseudo code implementation of the *flexible data set* utilised throughout the report see Appendix D. For each analysis after the *flexible data set* has been constructed, all companies for which the needed ratios exist are included. Winsorizations are performed, as described in the theory section, on the considered subset of companies. If all subsets of ratios would be considered, then $2^{63} - 1$ winsorizations would unavoidably have to be performed.

4.1.2 Potential Model Building Bias

A potential systematic bias could be introduced as, although it is known that the original data set contains a set of non-credit event companies agreeing well with the credit event companies there is no systematic implemented way to control each of the *flexible data set*-subsets. There does potentially exist combinations of ratios which have very few corresponding companies, but among which a model performs well. This gives reason for caution, but should, since the problem has been identified, be easy to avoid if the effect is deemed to have significant importance. Furthermore, the effect will be supervised by simply noting how many companies that are available for each analysis step.

4.1.3 Univariate Analysis

Because of the large number of covariates three initial tests are conducted for each ratio. The tests are Welch's t-test, Kolmogorov-Smirnov 2-sample test and a simple logistic regression. A Bonferroni correction will also be applied to the significance levels.

Welch's t-Test

The justification for choosing Welch's t-test in favor of the usual t-test is that it was confirmed by Levene's and Brown Forsythe's tests, that the ratios in many cases are rejected to be of equal variance. Testing for significantly different means is not imperative for application of logistic regression, but it is nevertheless an indication of some discriminative power.

Two-Sample Kolmogorov-Smirnov Test

Welch's t-test assumes normality, which is violated for many ratios considered in this report. A non-parametric test with few underlying assumptions is therefore warranted. Within this subset of tests the Kolmogorov-Smirnov 2-sample test is chosen. This test is hoped to capture differences in the ratios' behaviour beyond that which is discernible in means of artificially assigned parametric distributions.

Logistic Regression

Another important factor for determining if a ratio will have high discriminative power in the resulting model or not, is to consider it in a univariate regression. A simple logistic regression is performed for each of the ratios and it is noted whether the corresponding regression coefficient is significant on the 20 % level, based on the assumed univariate distribution of the regression coefficient, i.e. based on the Wald Statistic, where under H_0 the coefficient follows a normal distribution around 0. The 20 % significance level is chosen as a ratio may only be significant when regressed together with specific combinations of other ratios. Therefore, a lower restriction on significance than what may otherwise be appropriate is applied. A false positive is not disheartening here, but too many false negatives can quickly reduce the expected discriminative power of the resulting model.

Covariate Families

Based on the results from Welch's t-test at the 5 % significance level, the Kolmogorov-Smirnov 2-sample test at the 5 % significance level and the univariate regression at the 20 % significance level, the ratios are divided into four groups of ranked importance. The first, second and third and fourth group will consist of ratios for which all, two, one and zero tests are significant. The criteria are chosen, and deemed adequate, because they all, in distinct ways, indicate discriminative power; one test in terms of mean differences, one test in terms of distribution differences and the final test indicates discriminative power in a logistic regression model.

4.1.4 Correlation & Visual Analysis

Going beyond the univariate analysis, the correlations between the ratios are analysed. Keep in mind that the *flexible data set* is used throughout this correlation analysis, i.e. for each ratio-pair all the available companies in the estimation set, for that specific pair, are utilised. In situations where there is *very high* correlation (defined as having absolute value above 0.8) one of the ratios is deleted. If there is a difference in rank of the ratios with very high correlation, the family rank of the ratios decides which one to delete from

further analysis. In the cases where the ratios are of equal family-rank then the deletion is based on economic intuition and visual analysis of scatter plots.

For ratio pair-correlations classified as *high but not very high* (defined as having absolute value between 0.7 and 0.8) one of the following actions is performed, (1) Both ratios are kept; (2) An interaction term is formulated; or (3) One of the ratios is removed. The decision of which action to perform is based on, (i) Visual analysis of a 2-dimensional scatter plot; and (ii) Economic intuition. Absolute correlation above 0.7 is, for convenient notation, defined as *high or very high*. To more directly address point (2), if evident clustering of credit event companies or non-credit event companies in two dimensions emerge then the methodology is to try to formulate interaction terms, such as categorical terms or functions of ratios. However, one must tread carefully as one does not wish to overfit the estimation sample, through capturing noise which is unlikely to be present in the validation sample or another sample from the population. A covariate introduced by a correlation- or visual analysis procedure is included in the lowest ranked corresponding Covariate Family of the two or more ratios considered.

4.1.5 Controlled Selection of Covariates

Economic intuition will guide the inclusion and exclusion of additional covariates as needed and will also guide the exclusion of covariates if for example ratios are considered to be obvious linear combinations of other ratios.

4.1.6 Best Discriminant Stepwise Inclusion/Exclusion

In order to deal with the large number of covariates a stepwise algorithm approach was chosen. The Stepwise Inclusion/Exclusion utilised in this report is a custom built algorithm, constructed with the aim of maximizing discriminative power, while maintaining significance of the included covariates. At each step the algorithm adds the covariate with the best additional discriminative power, measured as the resulting model's sum of sensitivity and 1-specificity. After adding a covariate the p-values of all covariates, currently in the new model, are compared to 0.1.

This can be viewed as a two-step greedy algorithm. Step one consists of evaluating all neighboring models, in terms of discriminative power, and improving the model by including the best additional covariate. Step two consists of removing the covariate with the highest p-value if a p-value is above 0.1, which is also considered an improvement of the model. The next iteration only considers inclusion of additional variables not removed in the last iteration. The algorithm terminates if the model becomes too large or if the improvement is too small. The algorithm can therefore be considered greedy, since at each step the best possible move is being made until no

further improvements can be made. (Kleinberg & Tardos, 2005)

The algorithm is first applied to the covariates in Covariate Family 1, where the flexible data set has been applied with family-wide ratio restrictions. After the first run of the algorithm, a preliminary model is obtained. This model is the starting point for a second iteration of the algorithm where the covariates from Covariate Family 2 are considered. The flexible data set is in the second run restricted by use of the covariates in the preliminary model along with all covariates from Covariate Family 2. A third and a fourth run, including Covariate Family 3 and 4, is thereafter performed in a similar fashion.

See Appendix E for a pseudo code implementation of the Best Discriminant Stepwise Inclusion/Exclusion Algorithm along with application to the successive family expansion.

4.1.7 Final Ratios

The ratios that are retrieved after the fourth iteration of the algorithm will be the constituents of the final model. The model will be re-estimated, for illustrative purposes, with the *flexible data set* restricted only by the ratios included in the final model.

4.1.8 Correction for Finite Sample- & Rare Event Bias

Once the final ratios are found the coefficients are re-estimated, using the same data set, with the addition of the finite sample and rare event bias reduction techniques, as described in the theory. The resulting impact of the bias reductions are calculated and new levels of significance will be examined, and is expected to determine the choice of finite sample correction method. The impact of different population wide default rates is also investigated. The final model will be calibrated using one finite sample technique and rare event bias and a set of coefficients compensated for rare event bias and will be based on a suggested population-wide default rate. Significance of the ratios are evaluated using Wald Statistics, and in the case of Weighting, the Wald Statistics corresponding standard error is calculated using Huber-White (robust) standard errors as described briefly in the theory.

4.1.9 Uncertainty in Beta

As described in the theory, probabilities are estimated based on multiple draws of $\hat{\beta}$ in a discretised version of the integral in Equation 2.19. This is thus a Monte Carlo approach which is implemented with the use of *antithetic variates*. This is solely done to illustrate the impact that uncertainty in beta has, and is mostly of theoretical interest as it does not impact the point estimates of model coefficients, as it is excluded from cutoff optimizations.

4.2 Earlier Models

Altman's Z-score will be calculated for all companies in the validation data set that have all of Altman's variables available, in accordance with the application of the *flexible data set*. The discriminative power will be investigated by use of a classification table. Furthermore, the Altman's Z-score model will be recalibrated using the estimation data set. The re-calibrated model's discriminative power will be investigated by use of a classification table, and the model's performance will be compared to that of the original Z-score model. The same process is followed for Ohlson's O score.

4.3 Model Evaluation

Following the Model Building section, the resulting model will, as stated in the problem formulation, be evaluated against Altman's original Z-score and Ohlson's original O-score using CAP-curves and ROC-curves. Confidence intervals, along with unbiased estimates of the AUC and ROC, will for all models be calculated by Jackknifing.

A comparison of rank consistency between the model and the Nordic and US CDS and Rating data will also be presented. Both in plots based on actual ratings/rankings, in order to enable a visual guidance of the performance of the model, and through computation of correlations. More specifically Spearman's ρ will be calculated based on the rankings derived from the model and the Rating and CDS data sets respectively. For CDS spreads, as there are no ties in the CDS data, the calculations of Spearman's ρ along with its corresponding p-values are straightforward. However, as there are more companies than possible ratings it follows trivially, or by Dirichlet's principle, that some companies inevitably must have the same rating. If one does not fully trust the crude classification provided by the rating agencies, then it makes intuitive sense to measure the information truncation inherent in the agencies' ratings. This measure is obtained by resolving all tied ranks, in an unbiased way, which introduces variance. Confidence intervals are then based on a Monte Carlo scheme using the newly obtained ranks, called *pseudo ranks*. For each Monte Carlo simulation all ties within the data are resolved through randomly assigning an unbiased *pseudo ranking*, for each group of companies with the same rank, e.g. (1224), is resolved as (1234) or (1324) with 50 % probability each. Based on each *pseudo ranking* calculating Spearman's ρ is trivial, in extension the estimated expected value of Spearman's ρ follows by averaging over all calculated *pseudo correlations*. Confidence intervals for the true expected value will also be obtained, the width of the confidence interval measures the uncertainty introduced by the tied ranks.

5

Results

I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact.

– Winston S. Churchill

5.1 Model Building

5.1.1 Univariate Analysis

Following the construction of the estimation data set and after multiple recalibrations of the flexible data set some resulting descriptive statistics for all 63 ratios are presented in Appendix F. All of the ratios are tested by use of Welch’s t-test (at 5 % level), Kolmogorov-Smirnov’s two-sample test (at 5 % level) and a univariate logistic regression p-value test (at 20 % level). Depending on how many tests that are rejected for each ratio, the ratios are separated into one out of four Covariate Families as described in the Method.^aWhich tests that are rejected, together with the family classification for each ratio is presented in Appendix G.

5.1.2 Correlation & Visual Analysis

As part of the multivariate analysis 49 ratio pairs are identified for which the correlations are *high or very high*. The full correlation matrix is available in Appendix H, where a red field indicates *very high* correlation and an orange field indicates *high but not very high* correlation.^bThe 15 ratios that are removed due to having *very high* correlations are listed in Table 5.1 along

^aIt was attempted to apply a Bonferroni correction to the significance levels, but this resulted in too few rejected tests for convenient application of the covariate family approach.

with which ratio in the ratio pair that is kept and a short comment on the choice between the two.

#	Removed	In Favour of	Comment
2	Cash Ratio 1	QR	One of four ratios with pairwise high correlation
3	Cash Ratio 2	QR	
51	QA/TL	QR	
6	FCF/TL	CFO/TL	CFO/TL is used as Ohlson proxy
8	EBIT margin	EBITDA margin	EBITDA deemed more representative of payment means available
10	EBIT/TIntExp	EBITDA/TIntExp	
16	NI/TA	NI/MTA	One of three ratios with pairwise <i>high or very high</i> correlation
57	NI/MTA	NI/TL	
17	NI/TD	EBITDA/TD	One of four ratios with pairwise high correlation
33	TE/TD	EBITDA/TD	
55	Rev/TD	EBITDA/TD	
23	LTD/TD	STD/TD	Obviously, equal to -1
27	STD/TInvCap	WC/TA	Family Rank Difference
40	CL/TA	WC/TA	
52	QA/Rev	CA/Rev	CA/Rev exists for higher number of companies

Table 5.1: Ratios removed due to having *very high* correlation (>0.8)

The remaining 34 ratio-pairs are analysed in scatter plots. As most of this analysis merely shows noise and non-evident patterns, the description of this analysis is quite limited compared to the full amount of scatter plots analysed. More precisely plots are presented for two cases where decisions of inclusion/exclusion of ratios are made based on visual analysis and in Appendix I three examples representative for the full analysis are available. In all plots the top left and bottom right subplots show the univariate distribution of the ratios. The top right and bottom left plots show the bivariate distributions. In all plots red circles indicate credit event companies and blue circles indicate non-credit event companies. The difference in the bivariate distribution plots is the order the different colored circles are plotted and also which ratio goes on which axis.

In Figure 5.1 a combined scatter plot for WC/Rev and WC/TA is displayed. The bivariate distributions are deemed to be too similar to the univariate plots, i.e. too close to a linear relationship.^cTherefore, one of the ratios in the pair is removed. WC/TA has been used for removal of other ratios earlier, due to *very high* correlation. Those ratios that were removed earlier in the analysis do not have as high correlation to WC/Rev. So, if WC/TA is removed now, the formerly removed ratios would have to be rein-

^bLet us stress a point briefly mentioned in the method description. To fully understand and appreciate the correlation calculations, and in extension the correlation matrix, note that each point of the correlation matrix is calculated using a potentially very different set of companies, due to the flexible data set. The approach is believed to yield the best possible point-wise estimates for each of the pair-wise correlations within the correlation matrix, but as the companies vary across the rows and columns the correlations' inter-relationships should not be scrutinised, as contradictions may be present if interpreted as a normal correlation matrix. The matrix is furthermore not scalable to a classical estimate of the covariance matrix.

troduced. Instead, by removing WC/Rev, WC/TA represents an additional ratio for further analysis.

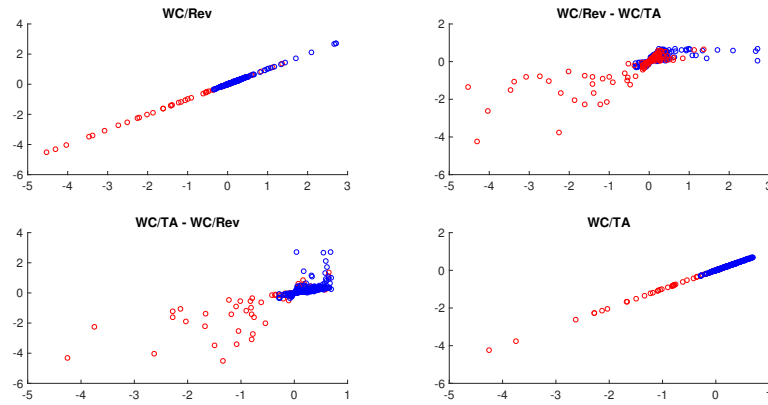


Figure 5.1: Visual analysis guidance for WC/Rev and WC/TA

In Figure 5.2, although no bivariate patterns are evident, the monotone discriminative power of WC/TA is identifiable. It also seems like WC/MCAP is a non-monotonic discriminant. The non-credit event companies seem to be located in the middle of the top left scatter-plot. This indicates that WC is either large or small in relation to MCAP for the credit event companies compared to the non-credit event companies. Therefore the log-odds of experiencing a credit-event is unlikely to be linear in the ratio WC/MCAP, which is an assumption for logistic regression. WC/MCAP is therefore removed in favor of WC/TA.

^cThe scatter plot is also seen to illustrate what is referred to as a self explanatory region as described in Appendix I.

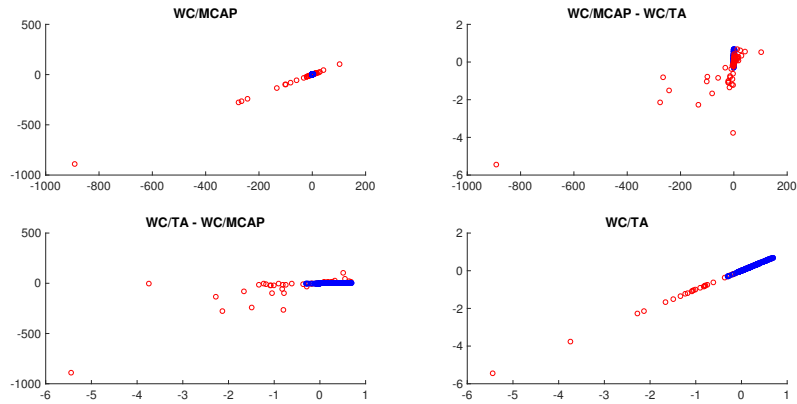


Figure 5.2: Visual analysis guidance for WC/MCAP and WC/TA

5.1.3 Controlled Selection of Covariates

Although not evident from the correlation analysis, there are some obvious duplicates or very similar ratios which need to be highlighted and removed before the model is built. The ratios that are removed due to this are found in Table 5.2.

Removed	In Favour of
Solvency	Solvency without Goodwill
EBITDA/Total Debt	EBITDA/Total Interest Expense
Net Income margin	EBITDA margin

Table 5.2: Ratios removed from controlled selection

Solvency - Solvency Without Goodwill

The ratios Solvency (with Goodwill) and Solvency without Goodwill intuitively share the same economic characteristics, they are equal if subtracting Goodwill from the nominator and denominator in the Solvency ratio. The company's Goodwill is thought to dilute the traditional solvency measure. Therefore Solvency (with Goodwill) is removed in favor of Solvency without Goodwill in hope that it will be a better indicator of the company's health.

EBITDA to Total Debt - EBITDA to Total Interest Expense

In the very high correlation analysis several variables are removed due to correlation with EBITDA/Total Debt. The conversion from Total Debt to Total Interest Expense is however company specific. The same amount of Total Debt may therefore accrue different amounts of Total Interest Expense for distinct companies. As EBITDA is a measure of how much money that

is available before interests are payed, it makes sense to compare against interest expenses. Considering also the definition of credit event along with the one year prediction horizon it is concluded that EBITDA to Total Interest Expense is a more accurate measure of risk than EBITDA to Total Debt.^d

Net Income margin - EBITDA margin

Both ratios are classified as profitability ratios. EBITDA is however a measure of profitability at an earlier stage in the Income Statement; namely, Before Interest, Taxes, Depreciation and Amortization. Cash generation is believed to be of higher importance than end of the line profits, for a one year credit event prediction. EBITDA is therefore believed to be a better measure to put in relation to sales.

Market Capitalization to Total Liabilities - Total Equity to Total Liabilities

It is noted in the correlation analysis that the ratio TE/TL has a very high correlation with MCAP/TL. Since MCAP does not exist for private companies it is thought to be suitable to make an exception and in this case not remove a ratio, as the method otherwise suggests. Instead, TE/TL and MCAP/TL are both kept in Covariate Family 1 for further analysis, but they will not be used simultaneously as they are thought to be too similar.^e

5.1.4 Final Covariate Families

In Table 5.3 the updated, and final, Covariate Families are displayed.

^dIn addition, CFO to Total Debt remains in Covariate Family 1 and CFO is more reasonable to use against Total Debt as CFO better captures how much money that is generated for repayments of principals.

^eIn the estimation set the inclusion of MCAP/TL as a ratio in Family 1 restricts the number of companies within the flexible data set, from 80 to 52 credit event companies and 160 to 155 non-credit event companies. Since the restriction is quite substantial among credit event companies, and since the rare events are more informative, as described in the theory section, the inclusion of MCAP/TL is questionable. One argument for including the ratio is the strength the covariate has in previous studies. (Altman, 1968)

Subset	#	Ratio
Covariate Family 1	1	Cash from Operations to Total Liabilities
	5	Cash to Total Assets
	9	EBIT to Total Assets
	11	EBITDA margin
	12	EBITDA to Net Debt
	14	EBITDA to Total Interest Expense
	19	Net Income to Total Liabilities
	20	Retained Earnings to Total Assets
	29	Solvency without Goodwill
	30	Total Equity to Long Term Debt
	32	Total Equity to Short Term Debt
	34	Total Equity to Total Liabilities
	37	log (Current Assets to Revenue)
	38	log (Current Assets to Total Assets)
	41	log (Quick Ratio)
	43	Working Capital to Total Assets
	53	log (Quick Assets to Total Assets)
	56	log (Market Capitalization to Total Liabilities)
	60	Change in Net Income (CHIN)
	63	Net Sales Change
Covariate Family 2	4	Cash to Revenue
	7	Interest Service Cover Ratio
	22	Long Term Debt to Total Assets
	24	Long Term Debt to Total Invested Capital
	25	OENEG
	26	Short Term Debt to Total Debt
	35	log (Total Liabilities to Total Assets)
	39	log (Current Liabilities to Current Assets)
	46	Accounts Receivable to Revenue
	58	Total Liabilities to Market Value Total Assets
	61	Current Asset Quality to Current Liability Quality
	62	INTWO
	Cov. Family 3	18
21		Intangibles to Total Equity
31		Total Equity to Net Debt
36		Size
49		log (Inventory Turnover)
50		Inventory Turnover to Working Capital
C. Fam. 4	44	Accounts Payable Turnover
	45	Accounts Receivable to Accounts Payable
	47	log (Accounts Receivable Turnover)
	48	Cash Conversion Cycle
	54	Revenue to Total Assets

Table 5.3: Updated and Final Covariate Families

5.1.5 Family-wise Stepwise Inclusion/Exclusion

Covariate Family 1

Analysis with MCAP/TL The stepwise algorithm is applied to Covariate Family 1 with MCAP/TL instead of TE/TL. The algorithm gives a model that only includes two ratios, MCAP/TL and NI/TL. The two ratios EBITDA/Rev and Cash/TA are temporarily considered by the algorithm, as the individual contribution to the discriminative power is high enough, but they are both excluded immediately due to too high p-values. Three models are therefore estimated; one model with the two main ratios (this model is called *MCAP-A* below), one model which in addition includes EBITDA/Rev (*MCAP-B*) and one model which replaces EBITDA/Rev with Cash/TA (*MCAP-C*). The coefficients and the p-values for each of the estimated models are presented in Table 5.4. The change of significance is explained by use of the *flexible data set*, which as usual allows more companies when less restrictive ratios are being considered. It turns out that for all three re-fitted models it is only MCAP/TL that is significant at the 5 % significance level. There are however reasons to doubt the applicability of MCAP/TL as foundation of a model. See Discussion 6.3.5 *Exclusion of Market Capitalization* for a more elaborate explanation. The ratio MCAP/TL, along with derivative models based on its inclusion, are excluded from further investigation.

	MCAP-A		MCAP-B		MCAP-C	
Coefficient	value	p-value	value	p-value	value	p-value
Intercept	-3.856	8.00E-15	-4.069	2.98E-12	-3.360	2.36E-10
Cash/TA	-	-	-	-	-8.756	0.0591
EBITDA/Rev	-	-	1.017	0.4354	-	-
$\log(\text{MCAP/TL})$	-2.133	9.79E-13	-2.173	1.68E-12	-2.101	4.65E-12
NI/TL	-1.469	0.1592	-1.746	0.1107	-1.582	0.1466

Table 5.4: Coefficient values and p-values of re-estimated stepwise models

Analysis without MCAP/TL If TE/TL is included in Covariate Family 1 and MCAP/TL is temporarily removed, then the custom stepwise algorithm yields the ratios presented in Table 5.5, when applied to Covariate Family 1.

The ratio Cash/TA is the last ratio to be added by the stepwise algorithm, but it is removed as it has a p-value of 0.1189, i.e. above 0.1. Due to the still relatively low p-value, inclusion of the ratio is worth to investigate further. Two models are therefore estimated; one model with the three main ratios in Table 5.5 (this model is called TE-A) and one model which in addition includes Cash/TA (TE-B). The coefficients and the p-values for each of the estimated models are presented in Table 5.6. Note that in the

re-estimation the ratio Cash/TA is significant. The change of significance is explained by the use of the *flexible data set*. Since all four ratio coefficients are now significant, and since inclusion of Cash/TA increase the discriminative power, the resulting model from the Family 1 stepwise section is the model consisting of the ratios in Table 5.5, with the addition of the ratio Cash/TA, i.e. this is the same as model TE-B in Table 5.6.

#	Description
10	Cash From Operations / Total Liabilities
18	EBITDA / Total Interest Expense
44	Total Equity / Total Liabilities

Table 5.5: Output from the stepwise algorithm when applied to Covariate Family 1 without MCAP/TL

Coefficient	TE-A		TE-B	
	value	p-value	value	p-value
Intercept	1.783	1.46E-10	2.192	8.48E-10
Cash/TA	-	-	-5.033	0.0396
CFO/TL	-8.220	8.50E-05	-7.999	1.50E-04
EBITDA/TIntExp	-0.248	1.01E-04	-0.271	1.37E-04
TE/TL	-3.681	5.72E-11	-3.760	8.43E-11

Table 5.6: Coefficient values and p-values of re-estimated stepwise models

Covariate Family 2

Starting with the Family 1-model as input, the stepwise algorithm is applied so that any of the ratios from Covariate Family 2 are allowed to be included, in addition to the four ratios already in the Family 1-model. The resulting model is the Family 1-model with the addition of ratio #26, Short-Term Debt to Total Debt. Table 5.7 shows the refitted coefficient values and p-values. Inclusion of the ratio STD/TD makes Cash/TA insignificant at the 5 % level. However, as described in the Method section, the algorithm is only allowed to remove ratios if the 10 % p-value level is breached. From a pragmatic standpoint, keeping the ratio must be considered, despite its slight insignificance, due to the ratio's economic soundness and intuitiveness. All five ratios are kept and this model is referred to as the *Family 1,2*-model.

Coefficient	Family 1,2-model	
	value	p-value
Intercept	1.869	4.60E-07
Cash/TA	-4.868	0.0530
CFO/TL	-8.437	7.07E-05
EBITDA/TIntExp	-0.294	3.06E-05
STD/TD	2.005	2.01E-03
TE/TL	-3.812	2.99E-10

Table 5.7: Coefficient values and p-values of re-estimated stepwise models

Covariate Families 3 & 4

Starting with the *Family 1,2*-model the stepwise algorithm does not suggest that any of the ratios in Covariate Family 3 or 4 should be included.

5.1.6 Final Ratio Model

The final model from the stepwise model building phase is equal to the *Family 1,2*-model as no ratios were added from Covariate Family 3 or 4, this model is referred to as the *Final Ratio Model*. Descriptive statistics for the final ratios are illustrated in Table 5.8. In Figure 5.3 a histogram of the *Final Ratio Model* output illustrates the discriminative power of the model and so does Classification Table 5.9. See Appendix J for an illustration of the univariate discriminative power for all of the ratios in the *Final Ratio Model*.

#	Ratio	Mean		Std.Dev.		Min		Median		Max	
		C	NC	C	NC	C	NC	C	NC	C	NC
1	Cash/TA	0.010	0.19	0.095	0.19	-0.49	-0.15	0.017	0.14	0.30	1.20
5	CFO/TL	0.068	0.11	0.075	0.11	0	0	0.044	0.064	0.31	0.42
14	EBITDA/TIntExp	0.95	16.36	1.73	47.66	-7.00	-29.04	1.06	6.08	7.20	367.82
26	STD/TD	0.30	0.13	0.42	0.21	0	0	0.032	0.036	1	1
34	TE/TL	-0.027	0.91	0.35	1.18	-0.65	-0.35	-0.083	0.66	1.74	9.59

Table 5.8: Descriptive Statistics for the final ratios, based on the 108 and 201 Credit Event Companies (C) and Non-Credit Event Companies (NC), in the estimation set

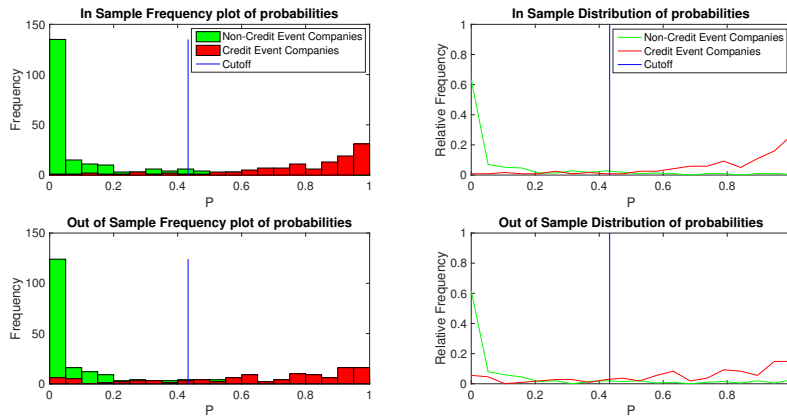


Figure 5.3: In- and out-of-sample frequency and distribution plots for the probabilities obtained from the *Final Ratio Model*

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	85 (TP)	29 (FP)	P. Pred. = 74.6 %
	Non-C. Event	23 (FN)	172 (TN)	N. Pred. = 88.2 %
		Sen. = 78.7 %	Spec. = 85.6 %	

Table 5.9: Out of Sample Classification Table for the *Final-Ratio Model*

5.1.7 Bias Correction

The finite-sample-, rare event bias-, and *uncertainty in beta* techniques, described in the theory section are applied to the final ratio-model. First, the Prior Correction and the Rare Event Bias Correction are jointly applied to the model, secondly the combination Weighting and Rare Event Bias Correction is considered and finally a crude analysis of the impact from *uncertainty in beta* is conducted. The true population-wide probability of default utilised for the bias correction techniques is 2.5 %.^f

Finite Sample & Rare Event Bias Correction

Table 5.10 presents the bias of the ratios along with the corresponding ratio coefficients and Wald statistics (bold if significant at 5 % level) when applying Prior Correction and Rare Event Bias Correction. The resulting coefficient values can be seen in the second column from the right. All the ratios except Cash/TA still have significant coefficients. Cash/TA is however, as previously, very close to being significant (compare Wald statistic

^fSlightly above the average of the issuer-weighted corporate default rate for 2008-2013. (Services, 2015)

to ± 1.96). The resulting model in Table 5.10 will be referred to as the *Bias Corrected Model*.

Table 5.11 contains the same type of information as Table 5.10, with the difference that Weighting instead of Prior Correction is applied. In contrast to Prior Correction, Weighting affects all of the estimated coefficients in the model. In Table 5.11 it is seen that combining Weighting and Rare Event Bias Reduction yields the same significant variables, based on Wald statistics adjusted with new custom standard errors, based on White-Huber (robust) standard error calculations.

In both Table 5.10 and Table 5.11 it can be readily seen that all coefficients have positive bias according to the Rare Event Bias Correction, drawing β towards $\mathbf{0}$ as expected.

Coefficient	β	PC: $\Delta\beta$	β_{PC}	RE: $\Delta\beta$	$\Delta\%$	β_{PC+RE}	Wald
Intercept	1.869	3.081	-1.212	0.142	7.6 %	-1.355	-3.720
CFO/TL	-8.437	0	-8.437	-0.840	10.0 %	-7.598	-3.643
Cash/TA	-4.868	0	-4.868	-0.147	3.0 %	-4.722	-1.911
EBITDA/TIntExp	0.294	0	-0.294	-0.304	11.7 %	-0.259	-3.746
STD/TL	2.005	0	2.005	0.307	15.3 %	1.697	2.663
TE/TL	-3.811	0	-3.811	-0.347	9.1 %	-3.464	-5.829

Table 5.10: Impact on ratios from Prior Correction (PC) and Rare Event Bias Correction (RE)

Coefficient	β	W: $\Delta\beta$	β_W	RE: $\Delta\beta$	$\Delta\%$	β_{W+RE}	Wald
Intercept	1.869	3.335	-1.466	0.112	6.0 %	-1.574	-15.948
CFO/TL	-8.437	1.381	-9.818	-0.988	11.7 %	-8.830	-2.152
Cash/TA	-4.868	-1.030	-3.838	-0.021	0.4 %	-3.817	-1.420
EBITDA/TIntExp	0.294	-0.194	-0.100	-0.013	4.3 %	-0.087	-68.460
STD/TL	2.005	0.178	1.827	0.229	11.4 %	1.598	2.522
TE/TL	-3.811	0.588	-4.400	-0.322	8.4 %	-4.078	-19.197

Table 5.11: Impact on ratios from Weighting (W) and Rare Event Bias Correction (RE)

Uncertainty in Beta

In order to illustrate the uncertainty in beta for the bias corrected model in Table 5.10, the probabilities are corrected by a Monte Carlo scheme. By the use of Antithetic Variates 1,000 versions of $\beta_{unbiased}$ are drawn from the asymptotic distribution $\beta_{unbiased} \sim \mathcal{N}(\beta, \text{Var}(\beta_{unbiased}))$, from which, by the consistency of $\beta_{unbiased}$, probabilities are calculated. Figure 5.4 shows the difference between the new estimated probabilities of experiencing a credit event, without uncertainty in beta, and the probabilities obtained by use of the point estimate of $\beta_{unbiased}$. The top part of Figure 5.4 illustrates

the actual differences and the bottom part illustrates the absolute differences. A horizontal line is drawn at 0 in the top plot, and vertical lines are drawn in both plots to separate credit event companies (red) from non-credit event companies (green). The average percentage point correction is 0.54 for credit event companies and 0.22 for non-credit event companies. The impact of *uncertainty in beta* is in many cases greater than the mean correction, which is especially true for the credit event companies, but the largest differences are, somewhat surprisingly, found among the non-credit event companies.

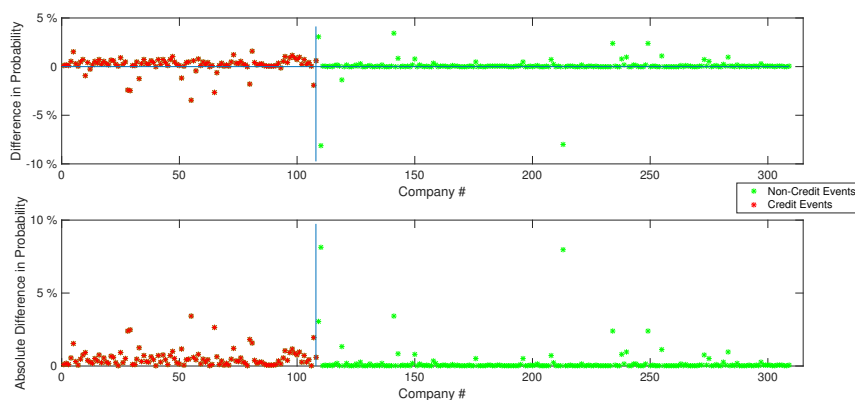


Figure 5.4: Illustrates the *uncertainty in beta* remaining in the bias corrected model from

Model Corresponding to Different Population-Wide Default Rates

Varying the population-wide default rate, τ , in the bias reduction step, changes the coefficients and the cutoffs of the models. The bias-reduction techniques used for this section are Prior Correction and Rare-Event Bias Correction. The following analysis contains variations of the *Bias Corrected Model* referred to above, which in that case used $\tau = 2.5\%$ as population-wide default rate. Table 5.12 shows the model coefficients and the optimal cutoff for different population-wide default rates τ .

τ	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Intercept	-3.026	-2.345	-1.909	-1.596	-1.355	-1.159	-0.994	-0.851	-0.725	-0.611
CFO/TL	-5.923	-6.482	-7.028	-7.379	-7.598	-7.738	-7.831	-7.894	-7.983	-7.969
Cash/TA	-5.367	-4.995	-4.853	-4.774	-4.722	-4.684	-4.656	-4.635	-4.619	-4.607
EBITDA/TIntExp	-0.252	-0.250	-0.254	-0.257	-0.259	-0.261	-0.262	-0.264	-0.265	-0.265
STD/TD	1.204	1.494	1.598	1.657	1.697	1.727	1.751	1.769	1.785	1.798
TE/TL	-2.989	-3.169	-3.312	-3.405	-3.464	-3.505	-3.534	-3.555	-3.571	-3.584
Cutoff	0.83 %	1.69 %	2.49 %	3.18 %	3.87%	4.60 %	5.34 %	6.07%	6.80 %	7.52 %

Table 5.12: Models corresponding to different population-wide default rates $\tau = 0.5\%, \dots, 5\%$

5.1.8 5-Factor Model

The final models that are obtained all use the five ratios from the *Final Ratio Model*. The coefficients of these ratios, and the cutoff, depend on the chosen population-wide default rate as shown in Table 5.12. In the model evaluation section below the model corresponding to 2.5 % default rate will be used. This model is hereafter called the *5-Factor Model* and is restated for convenience:

$$L = -1.355 - 7.598X_1 - 4.722X_2 - 0.259X_3 + 1.697X_4 - 3.464X_5 \quad (5.1)$$

$$p = \frac{1}{1 + e^{-L}} \quad (5.2)$$

Where,

$$X_1 = \text{CFO/TL}$$

$$X_2 = \text{Cash/TA}$$

$$X_3 = \text{EBITDA/TL}$$

$$X_4 = \text{STD/TD}$$

$$X_5 = \text{TE/TL}$$

$$\text{cutoff} = 3.87 \%$$

The performance of the *5-Factor Model* is depicted in Figure 5.5 and Table 5.13. The discriminative power is barely visible in the histogram plot as many probabilities are pulled towards the anticipated population-wide default rate, $\tau = 2.5\%$. By comparing the classification tables for the *5-Factor Model* and the *Final Ratio Model* in Table 5.13 and Table 5.9 it is seen that the performance is almost unchanged after applying bias corrections, as the number of miss-classifications only differ by two companies.

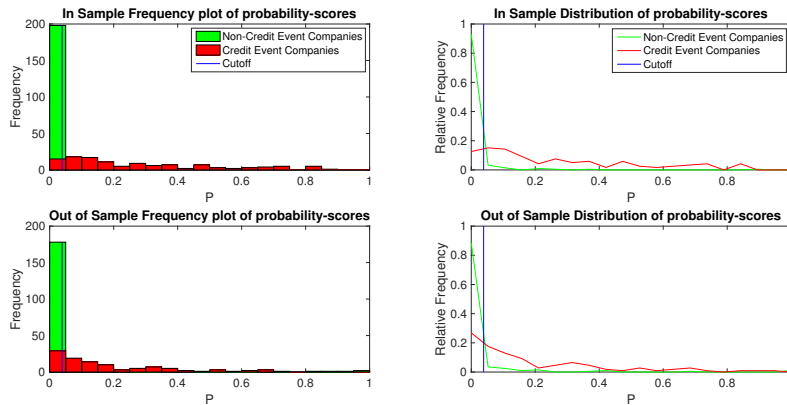


Figure 5.5: In- and out-of-sample frequency and distribution plots for the probabilities obtained from the *5-Factor Model*

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	84 (TP)	28 (FP)	P. Pred. = 75.0 %
	Non-C. Event	24 (FN)	173 (TN)	N. Pred. = 87.8 %
		Sen. = 77.8 %	Spec. = 86.1 %	

Table 5.13: Out of Sample Classification Table for the *5-Factor Model*

5.2 Earlier Models

In this section the performances of Altman's and Ohlson's original models are presented. The original models are also re-estimated with the estimation set by use of logistic regression and new coefficients and cutoffs are presented. The discriminative power of the original and re-estimated models are also compared in the validation set.

5.2.1 Altman Z-Score

Original

In the validation set there are 58 credit event companies and 224 non-credit event companies that have information for all of Altman's ratios. These companies are assigned a Z-score value according to Altman's model. Companies are classified in the three groups depending on their Z-score; high bankruptcy risk, grey zone and low bankruptcy risk. The discriminative ability for Altman's model is illustrated in Figure 5.6 and Table 5.14. The plots to the left in Figure 5.6 show the amount of companies in the different groups and the plots to the right in the figure show a more granular distribution of the companies.

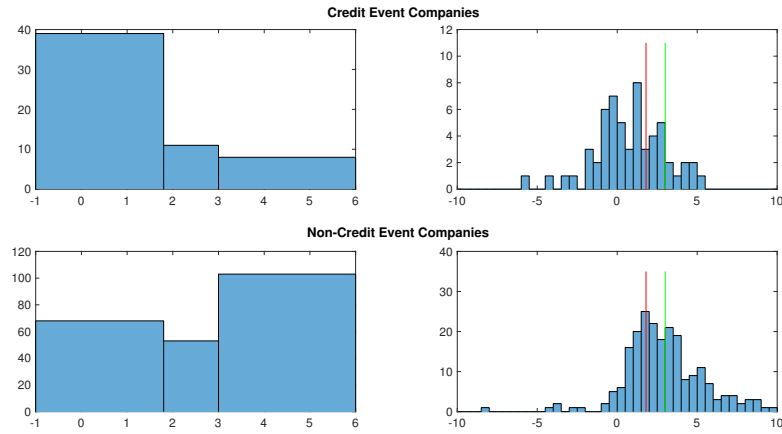


Figure 5.6: Distribution of Z-scores for credit and non-credit event companies

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	39 (TP)	68 (FP)	P. Pred. = 36.4 %
	Grey Zone	11	53	
	Non-C. Event	8 (FN)	103 (TN)	N. Pred. = 92.8 %
		-	-	

Table 5.14: Discriminative ability of Altman's original model

Re-Estimated

In the estimation set there are 73 credit event companies and 224 non-credit event companies that have information for all of Altman's ratios. These companies are used to re-estimate the coefficients for Altman's ratios. The re-estimated coefficients and their respective p-values are found in Table 5.15. Only one of the ratios from Altman's model has a p-value above 10 %. This is an indication that Altman's original ratios still have good predictive power, note especially MCAP/TL which has a p-value of $5E-7$. Table 5.16 illustrates the out-of-sample discriminative ability of the re-estimated version of Altman's model in a Classification Table. See Appendix K for in- and out of sample frequency and distribution plots of the probabilities obtained from Altman's re-estimated model.

	Altman Re-Estimated	
Coefficient	value	p-value
Intercept	1.1669	0.0166
WC/TA	-4.868	0.0284
RE/TA	-8.437	0.0817
EBIT/TA	-0.294	0.0993
MCAP/TL	2.005	5.06E-07
Rev/TA	0.1958	0.3369
Cutoff	40.1 %	-

Table 5.15: Altman's re-estimated coefficient values and p-values

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	41 (TP)	25 (FP)	P. Pred. = 62.1 %
	Non-C. Event	17 (FN)	199 (TN)	N. Pred. = 92.1 %
		Sen. = 70.7 %	Spec. = 88.8 %	

Table 5.16: Out of Sample Classification table for the re-estimation of Altman's original model

5.2.2 Ohlson's O-Score

Original

In the validation set there are 110 credit event companies and 231 non-credit event companies that have information for all of Ohlson's ratios. Companies are assigned an O-score value according to Ohlson's model. The corresponding probabilities are then compared to the cutoff probability of 0.5. If they are above, they are considered as having high risks of bankruptcy, otherwise as having low risks of bankruptcy. The discriminative ability for Ohlson's model is illustrated in Figure 5.7 and Table 5.17. It is evident from Figure 5.7 that Ohlson's model is better at identifying credit event companies than non-credit event companies.

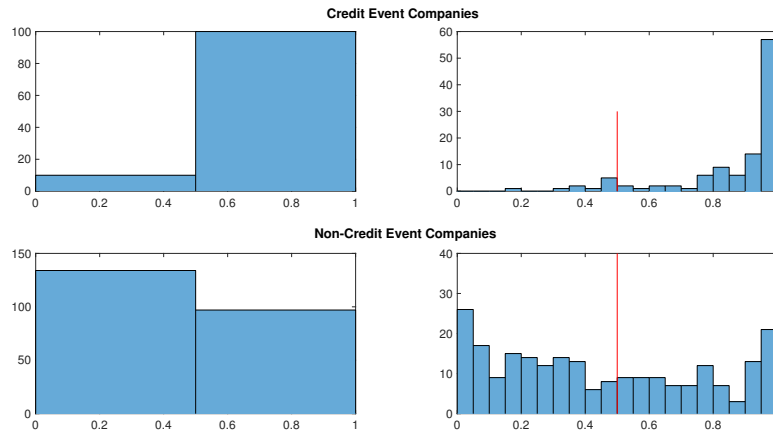


Figure 5.7: Distribution of O-scores for credit and non-credit event companies

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	100 (TP)	97 (FP)	P. Pred. = 50.8 %
	Non-C. Event	10 (FN)	134 (TN)	N. Pred. = 93.1 %
		Sen. = 90.9 %	Spec. = 58.0 %	

Table 5.17: Out of Sample Classification table for Ohlson's original model

Re-Estimated

In the estimation set there are 125 credit event companies and 237 non-credit event companies that have information for all of Ohlson's ratios. These companies are used to re-estimate the coefficients for Ohlson's ratios. The old coefficients, the re-estimated coefficients and the respective p-values for all re-estimated coefficients are found in Table 5.18. Only five out of Ohlson's nine ratios are significant in the re-estimation. Table 5.19 illustrates the out-of-sample discriminative ability of the re-estimated version of Ohlson's model in a Classification Table. The re-estimated model classifies 90.0 % of the non-credit event companies correctly, compared to just 58.0% for Ohlson's original model. See Appendix K for in- and out of sample frequency and distribution plots of the probabilities obtained from Ohlson's re-estimated model.

	Ohlson O-Score	Ohlson Re-Estimated	
Coefficient	old value	new value	p-value
Intercept	-1.320	0.035	0.980
Size	-0.407	-0.241	0.172
TL/TA	6.030	0.3270	0.260
WC/TA	-1.430	-2.193	0.0360
CL/CA	0.0757	0.2165	0.386
NI/TA	-2.370	-2.741	0.0298
CFO/TL	-1.830	-8.900	1.73E-05
INTWO	0.285	1.276	1.68E-03
OENEG	-1.720	0.903	0.0420
CHIN	-0.521	-0.280	0.402
Cutoff	50.0 %	39.9 %	-

Table 5.18: Illustrating the coefficient values (new and old) and the p-value. Bold values are significant at 5 % level

		Outcome		
		C. Event	Non-C. Event	
Prediction	C. Event	83 (TP)	23 (FP)	P. Pred. = 78.3 %
	Non-C. Event	27 (FN)	208 (TN)	N. Pred. = 88.5 %
		Sen. = 75.5 %	Spec. = 90.0 %	

Table 5.19: Out of Sample Classification table for Ohlson’s re-estimated model

5.3 Model Evaluation

In this section the *5-Factor Model* is compared to Altman’s and Ohlson’s models by use of the Cumulative Accuracy Profile (CAP) and Receiver Operating Characteristic (ROC) curves. All the model evaluation tools are based on the validation data set, the evaluation is thus on *unseen*, and of course non-winsorised, data. Due to the *flexible data set*, the curves are based on non-identical subsets of the validation data set. Altman’s model uses 58 credit event companies and 224 non-credit event companies, Ohlson’s model uses 110 credit event companies and 231 non-credit event companies and finally the *5-Factor Model* uses 108 credit event companies and 201 non-credit event companies.

5.3.1 Cumulative Accuracy Profile & Area Under Curve

In Figure 5.8 the CAP is plotted for each model. Visually, the *5-Factor Model* appears to outperform both Ohlson’s and Altman’s models. But the data set differences introduce a problem for visual comparison of the CAP

curves, as the ideal shape for each curve is different. The AUC-statistic presented in Table 5.20 normalises performance across sample-differences, and provide a fair comparison for the different models. All three models have AUC-scores between 0.5 and 0.9 but the *5-Factor Model* is performing significantly better, as seen by the Jackknife confidence intervals in Table 5.20. Note also the wide confidence interval for Altman’s model indicating high sensitivity to the sample changes within the Jackknife framework. Based on the confidence intervals the *5-Factor Model* is likely to be superior to Ohlson’s model which in turn is likely to be superior to Altman’s model.

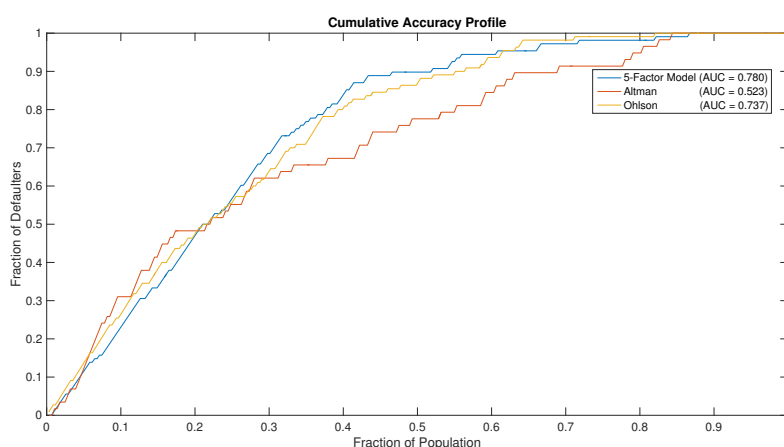


Figure 5.8: Cumulative Accuracy Profile for Ohlson’s & Altman’s models and the *5-Factor Model*

	<i>5-Factor Model</i>	<i>Altman’s Model</i>	<i>Ohlson’s Model</i>
97.5 %	0.784	0.701	0.742
Estimate	0.780	0.546	0.737
2.5 %	0.775	0.390	0.733

Table 5.20: Jackknife Confidence Intervals of AUC for the three models

5.3.2 Receiver Operating Characteristic

The ROC curves for each of the models are depicted in Figure 5.9, among which the *5-Factor Model*’s curve has the most ideal shape. Initially the models have roughly equal discriminative power but the strength of the *5-Factor Model* is evident after the first 40 % of correctly classified credit event companies (Y-axis ≈ 0.4). The superiority of the *5-Factor Model* is furthermore confirmed by the ROC. The ROC-statistic is for the *5-Factor Model* and Ohlson’s model close to 0.9 (0.890), which by Hosmer et al. (2013) indicates excellent, and close to outstanding performance. The same value

for Altman’s models have acceptable discriminative power according to the same authors. The Jackknife confidence intervals for the ROCs, provided in Table 5.21, show that the *5-Factor Model* is likely to be better than Ohlson’s model, which in turn is likely to be better than Altman’s model.

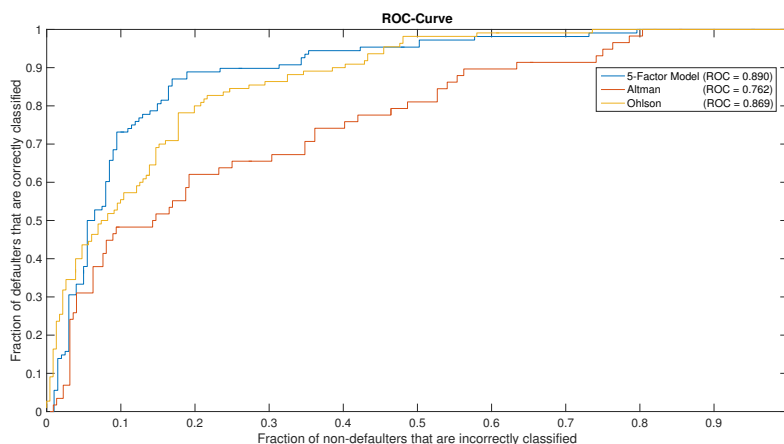


Figure 5.9: Receiver Operating Characteristic for Ohlson’s & Altman’s models and the *5-Factor Model*

	<i>5-Factor Model</i>	<i>Altman’s Model</i>	<i>Ohlson’s Model</i>
97.5 %	0.892	0.766	0.871
Estimate	0.890	0.762	0.869
2.5 %	0.888	0.757	0.867

Table 5.21: Jackknife Confidence Intervals of ROC for the three models

5.3.3 Credit Rating Data Comparison

In the US rating data set 159 companies are missing at least one required ratio for use of the *5-Factor Model*, these companies were thus all excluded. The final US rating data set contains 895 companies. In the Nordic data set all 35 companies have all the ratios for the *5-Factor Model*.

Figure 5.10 and Figure 5.11 displays all US and Nordic company’s rankings based on the *5-Factor Model* against the company’s corresponding S&P rating. The default probabilities are ranked from low (left) to high (right) on the X-axis. On the Y-axis the ratings are depicted in descending order of quality. To put it more clearly, the lower left corner should have the best ranked companies according to the *5-Factor Model* and the best ratings according to S&P.

For both the US and Nordic data sets a positive monotone relationship is evident, which indicates that the the rankings can be used as a proxy

for S&P's ratings. The top right and bottom left boxed areas of Figure 5.10 show that the *5-Factor Model* manages to capture the worst and best rated companies respectively within the US data, with only few errors.^g One clearly notes that the *5-Factor Model* has difficulties for the US companies in the intermediate ratings, approximately from B to A, see Discussion 6.2.3 for a plausible explanation. In Figure 5.12 the median ranking for each rating category is illustrated through the inclusion of a blue line.^h Due to the small sample size of the Nordic dataset extensive inferences based on visual analysis in Figure 5.11 is difficult.

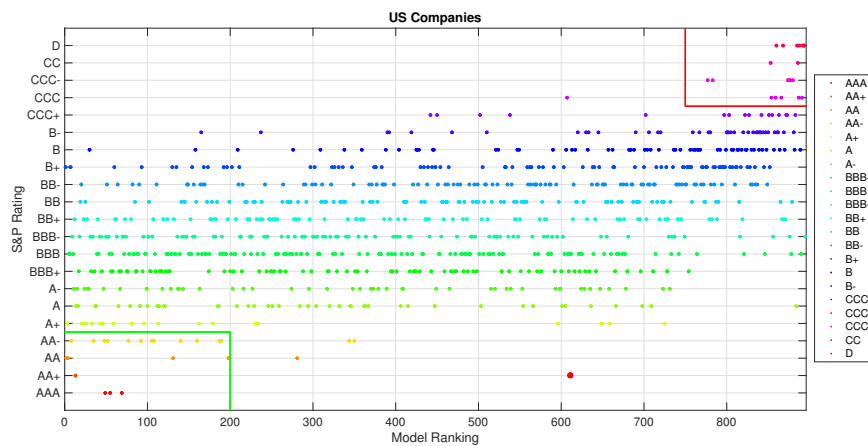


Figure 5.10: Rankings from the *5-Factor Model* displayed against S&P's corresponding rating for US companies

^gNote the marked outlier at coordinates (611,AA+), in the bottom right of Figure 5.10. Curious about the obvious mismatch we decided to look further at this company, which happened to be General Electric Company. The *5-Factor Model* rank measures the company's relative health as of its EOY 2015. We present two facts in this footnote: (1) In roughly the same period, December 7 2015, GE terminated an agreement with Electrolux from which they received \$175 m in a breakup fee from a sale of its Appliances business, the sell of the Appliances business hastily went into agreement with Haier for \$5.4 bn on 15th January 2016. (2) GE experienced a drop in Net Income from \$15.2 bn in 2014 to negative \$6.2 bn in 2015, and EBITDA went from \$15.5 bn in 2014 to \$12.0 bn in 2015. We leave it to the reader to draw conclusions whether these facts indicates a desperate need or excess of capital within GE as of EOY 2015.

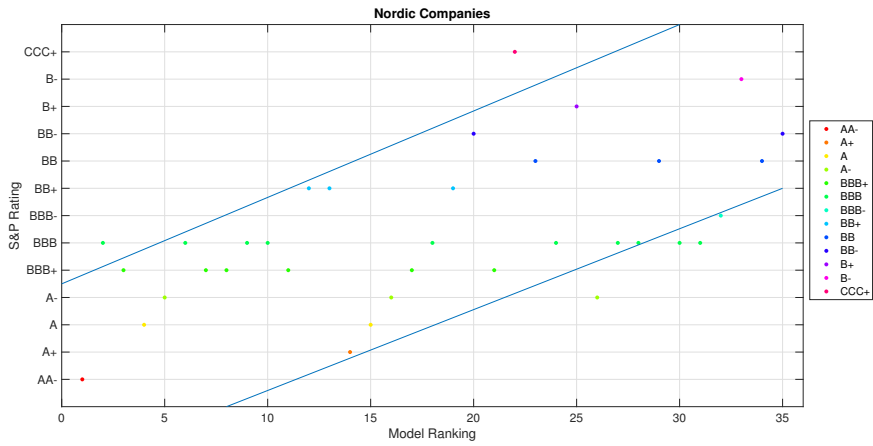


Figure 5.11: Rankings from the *5-Factor Model* displayed against S&P's corresponding rating for Nordic companies

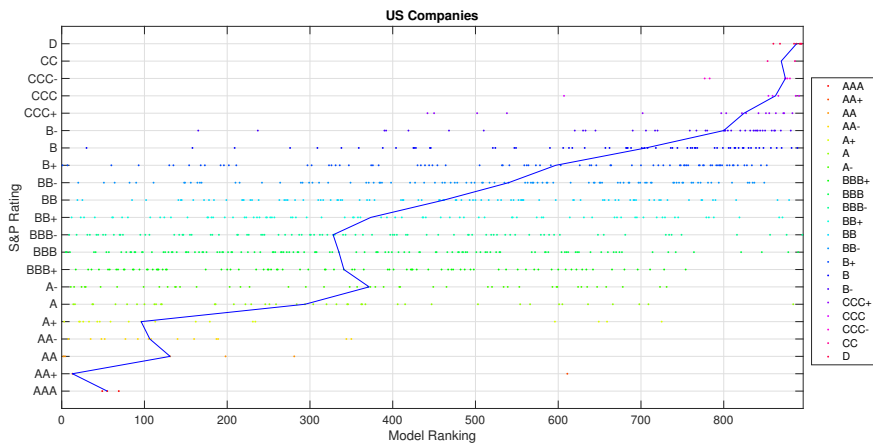


Figure 5.12: Rankings from the *5-Factor Model* displayed against S&P's corresponding rating for US companies with median ranking marked for each set of companies with distinct ratings

^hAlthough the median is a relatively robust statistic, and thus insensitive to outliers, the outlier located at coordinates (611,AA+) was excluded from the construction of the median line. This is because it would greatly have distorted the appearance of the median line, as there are only two AA+ rated companies.

ρ	<i>US</i>	<i>Nordic</i>
97.5 %	0.516	0.626
Estimate	0.511	0.567
2.5 %	0.506	0.506

Table 5.22: Monte Carlo based Confidence Intervals for Spearman’s coefficient between rankings from the *5-Factor Model* and S&P’s corresponding rating for the US and Nordic markets

5.3.4 CDS Data Comparison

The comparison of the ranking based on the CDS spreads and the ranking based on the *5-Factor Model*, relies on visual analysis and Spearman’s rank coefficient. For 97 out of the 115 US firms and 23 out of the 29 Nordic firms, all required ratios are available.

In Figure 5.13 the CDS rank is displayed against the *5-Factor Model* rank. A low rank on the X and Y axis indicates a low spread and a low probability of default respectively. For the US companies there appears to be a positive correlation. Unfortunately it is not possible to reach the same conclusion for the Nordic companies. The red and blue lines are drawn for the reader’s convenience. The red lines are drawn to indicate that the highest spreads and probabilities are grouped together. The blue lines are drawn to indicate the monotonic correlation that is visible. For the highest spreads the model assigns the highest probabilities and most observations are almost rank-consistent! The Spearman’s rank correlation coefficients and the corresponding p-values are shown in Table 5.23. Since the CDS spreads are without ties it is possible to analytically compute p-values, which should be interpreted as the probability of having at least the correlation presented in Table 5.23.

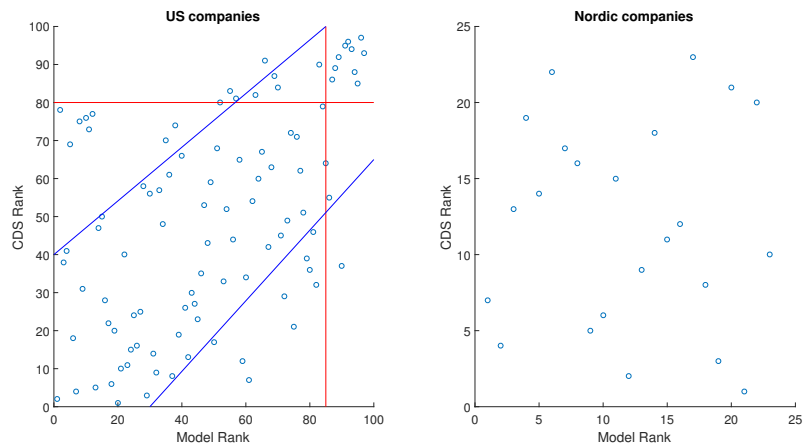


Figure 5.13: Rankings from the *5-Factor Model* displayed against the corresponding CDS spreads' rankings, for US and Nordic companies

ρ	<i>US</i>	<i>Nordic</i>
Estimate	0.501	-9.90E-03
p-value	2.64E-07	0.966

Table 5.23: Spearman's ρ , with corresponding p-values, based on ranks of the *5-Factor Model* output and CDS spreads, for US and Nordic companies

6

Discussion

Prediction is very difficult, especially about the future.

– Niels Bohr

In this chapter, (1) The resulting model from the model building stage, i.e. the *5-Factor Model*, is presented and discussed in detail; (2) All models' performances are discussed and for the *5-Factor Model* the performance is related to CDS and Rating data; (3) Emerging issues from the model building stage are discussed, along with other important clarifications, concerning for example Bias Correction; (4) Examples are presented concerning how to apply the *5-Factor Model* in practice.

6.1 The 5-Factor Model

The discussion below concerns the *5-Factor Model* (restated below for convenience) along with intuitive interpretation, adequacy and reasonability of its five ratios.

$$L = -1.355 - 7.598X_1 - 4.722X_2 - 0.259X_3 + 1.697X_4 - 3.464X_5 \quad (6.1)$$

$$p = \frac{1}{1 + e^{-L}} \quad (6.2)$$

Where,

$$X_1 = \text{CFO/TL}$$

$$X_2 = \text{Cash/TA}$$

$$X_3 = \text{EBITDA/TL}$$

$$X_4 = \text{STD/TD}$$

$$X_5 = \text{TE/TL}$$

$$\text{cutoff} = 3.87 \%$$

6.1.1 Discriminative Function

As a default option the proposed model is the *5-Factor Model* calibrated with $\tau = 2.5\%$ as population wide default rate. Changing τ mainly impacts the intercept and the cutoff, the models are therefore expected to perform similarly. However, the 2.5% model is the only model with performance verified in this report, especially with respect to rating and CDS rank consistency. If a different population wide default rate is anticipated, then the appropriate coefficients in Table 5.12 should be retrieved. For ranking purposes however, as the rankings are expected to be unaffected by the choice of τ , the default option should be to use $\tau = 2.5\%$. It is only when actual probability estimates are of interest that one should resort to Table 5.12, a more elaborate discussion follows in 6.4.

6.1.2 Intuitive Explanation of Ratios

Below the five ratios are considered in isolation. For the reader's convenience the ratios are individually conceptually captured by a short *factor name*. All factors capture distinct parts of a company's financial health. The factor names are: *Cash Generation*, *Cash Cushion*, *Interest Coverage*, *Maturity Structure* and *(Reverse) Leverage Position*.

Cash From Operations/Total Liabilities - Cash Generation

CFO/TL measures a company's long term ability to generate capital for its liabilities through its yearly cash from operations. A negative coefficient is thus to be expected as an increase in the ratio indicates that a company generates more cash in relation to its liabilities, and should therefore have better capability to eventually meet those liabilities. Note that this is the proxy for a ratio originally considered in Ohlson (1980), namely Funds Provided by Operations / Total Liabilities.

Cash/Total Assets - Cash Cushion

Cash/TA gives a normalised measure of the amount of cash the company has available. The corresponding coefficient has a negative sign as a company with higher cash reserves is more likely to be able to meet its contractual obligations. That the corresponding coefficient is less significant than the average coefficient within the *5-Factor Model* is intuitive as the measure fluctuates throughout the year, and it is furthermore not clear that a company desires high cash deposits as it can indicate inefficiency. An optimal composition of Cash to Total Assets is also expected to be industry dependent.

EBITDA/Total Interest Expense - Interest Coverage

EBITDA/Total Interest Expense, also known as the *EBITDA to Interest Coverage Ratio*, is a ratio measuring how many times a company could pay of its interest expenses with its available earnings. If the ratio is high, then a company is theoretically capable of repaying its interest expenses several times; meanwhile, if it is low then the interest expenses heavily burden the bottom line financial result. In effect this measure gives a normalised Interest Rate cost in relation to the true earnings of the company, naturally taken before interest expenses. Having a ratio greater than 1 is essential as otherwise the company is forced to use cash to pay off its interest expenses. The negative coefficient for EBITDA/TIntExp indicates that an increase of the ratio decreases the probability of default.

Short-Term Debt/Total Debt - Maturity Structure

If this ratio is high, then the company has a large degree of debt to pay in the near future. The positive coefficient is therefore intuitive, as an increase in the ratio indicates a greater need for short-term capital, and in extension bad health. However, the ratio does not take into account how much total debt the company has. If the amount of total debt is very low, and the debt is mostly short-term, then the impact on the probability could be unjustly high, since the amount of total debt could be minuscule in relation to the size of the company.

Total Equity/Total Liabilities - (Reverse) Leverage Position

Total Equity to Total Liabilities is the reciprocal of the commonly known Debt/Equity ratio. The reciprocal is chosen in this report as shifting signs are considered less intuitive for denominators than for nominators. As a rule of thumb, the more a company relies on liabilities to finance its operations the riskier it is. The negative coefficient for the ratio in the model indicates that an increased *leverage* gives a higher probability of default. The ratio is not perfect however, as companies within distinct industries often require different amounts of leverage to stay competitive. A high equity/liability ratio may be common practice in one industry, but for another industry the same amount of leverage may be undesired. The *5-Factor Model* makes no such note of industry standards.

6.1.3 Variable Splitting

Some of the ratios are, after careful consideration, not straightforward to interpret. The issue can occur when the numerator in the ratio changes sign from positive to negative. This is the case for three of the ratios above, namely; CFO/TL, EBITDA/TIntExp and TE/TL. For example, if EBITDA

is positive, then an increase in the ratio is considered beneficial for the company as more funds are available for the payment of interest expenses. But if EBITDA is less than zero, complications arise. A negative EBITDA is an indicator of bad health, but coupled with a large interest expense the impact of the negative EBITDA is mitigated. The model interprets a negative EBITDA coupled with high interest expenses as better than the same EBITDA with low interest expenses, which naturally should not be the case.^a

The complications arising from the possibility of negative numerators can be mitigated by splitting the ratios into two variables, one which is only active when the numerator is positive and one which is only active when the numerator is negative. This simple solution could improve the predictive ability of not only credit event models but also in a more general GLM framework whenever ratios with non-positive support are used as covariates.

6.2 The Performance of the Models

All in all, the performance of the models implemented and re-estimated is seen as remarkable! Careful consideration was put into the search of a credit event data set and a corresponding representative non-credit event data set. Ideally though, an alternative approach would be preferable; construct a population as Moody's full corporate coverage and then consider the two mutually exclusive and collectively exhaustive subsets, credit event companies and non-credit event companies. The method described in the Data section does not guarantee that all the companies present in the non-credit event data set would have appeared as credit event companies if a credit event would have occurred for them. But this problem is not unique for this report, and not feasible to resolve.

However, by testing the model with out-of-sample data confidence is gained that the *5-Factor Model* works well within the boundaries of the gathered data-set. Only time will tell if the model continues to perform well. The ratios in the *5-Factor Model*, as discussed above, make intuitive sense, which is believed to be very important.

6.2.1 Earlier Models

The uncalibrated Altman model correctly classifies 50.4 % (see Table 5.14) of the companies while the re-estimated model correctly classifies 85.1 %

^aTo further elaborate on the subject and possibly cause some confusion consider the following conversation: *A linguistics professor at MIT was lecturing his class the other day. "In English," he said, "a double negative forms a positive. However, in some languages, such as Russian, a double negative remains a negative. But there isn't a single language, not one, in which a double positive can express a negative." A voice from the back of the room piped up, "Yeah, right."*

(see Table 5.16). Ohlson's original model correctly classifies 68.3 % (see Table 5.17) of the companies in the validation set while the re-estimated version classifies 85.3 % correctly (see Table 5.19). Re-estimation of the older models yields better discriminative power in terms of overall correct classifications. Ohlson's original model is very good at identifying credit event companies, but as it also classifies a lot of non-credit event companies wrong, the positive predictive value is only 50.8 %. The low value displays that 49.2 % of the predicted credit events are false positives, a significant amount. The re-estimated version has a positive predictive value of 78.3 % which indicates that the model is more certain of its identified credit event companies.

The original calibration of Altman's model is based on Manufacturing companies and Ohlson's original model is based on Industrial companies. The data sample used in this report includes additional industries, which is expected to be disadvantageous for Altman's and Ohlson's original, and possibly re-estimated, models. That the models still show acceptable discriminative power 30-50 years after their development is however a remarkable feat! The difference between the original and re-estimated models are likely due to the different industry compositions and also due to change in business climate over time. Companies are likely structurally different today compared to 30-50 years ago. Altman's model is furthermore re-estimated using logistic regression rather than MDA, which could also impact the performance of the model adversely. There is also a difference in how credit events are defined in Altman (1968) and Ohlson (1980) compared to this thesis, which can have impacted the performances.

6.2.2 The 5-Factor Model

As seen from Result section 5.3 the CAP-curve, the AUC-score, the ROC-curve and the ROC-score of the *5-Factor Model* are all better than the older models' curves and scores. Considering the ROC-scores along with the fact that the O-Score has 4 more degrees of freedom, in terms of additional parameters to estimate, makes the *5-Factor Model* superior.

6.2.3 Credit Rating Analysis

As can be seen by Table 5.22 the estimated Spearman-coefficients are significantly non-zero, for both the US and Nordic markets. The estimates are both even greater than 0.5 which by Cohen's Standard indicates a strong positive monotone relationship between the *5-Factor Model* and the S&P ratings. To fully appreciate this feature of the model, consider that all these tests are performed on data drawn from new populations, and not different samples from the same populations used in the model building and model calibration steps. The high level of correlation, seen on both the US and

Nordic markets, is also impressive and shows that the model can be used for assessing the credit quality of both US and Nordic companies. Furthermore, a perfect rank correlation to the S&P rating data is not necessarily the ideal correlation, as it is not evident that the S&P ratings indicate the true credit quality.

The 95 % confidence intervals are relatively tight for the US companies, indicating that the information loss due to ties within the rating data can be seen as having limited impact. For the Nordic data the same conclusion is not as evident, which is likely due to the small sample size of Nordic companies.

Updating Frequencies of Ratings

Another factor possibly reducing the rank correlation to the S&P rating data is the lagged updating frequency inherent in the methodology utilised by the rating agencies. There is no set time in which a credit rating update will be given and the incentive structure of the rating agencies introduces a potential bias, as the rated companies can pay the agencies for an update when preferable. Meanwhile the accounting data used for the *5-Factor Model* is usually available quarterly and it is standardised. The identified regions of Figure 5.10 indicate that the *5-Factor Model*'s ranks agree especially well for the highest ranked companies, where only few large rank differences are present. This can potentially be explained by the lagged updating frequencies, as companies that are performing well are more inclined to pay for an updated rating. That the model agrees with the highest ranks can therefore be interpreted as an indication that the *5-Factor Model* is working. The model also seems to agree well for the lowest rated companies. We believe that this also could be due to more accurate ratings from the rating agencies for these companies. But in this case the updated ratings are likely to be initialised by the rating agencies, as they don't want inaccurate ratings they are likely to want to update the deteriorating companies' ratings. Conclusively, since the *5-Factor Model* seems to have higher rank consistency for the highest and the lowest rated companies, where one can argue that the ratings are the most accurate, and less biased, this can be seen as an indication that the *5-Factor Model* gives a more accurate credit quality assessment of companies than the actual ratings.

6.2.4 CDS Analysis

From the CDS analysis the *5-Factor Model* is somewhat able to rank the companies to the CDS spreads for the US market. A Spearman's ρ for the US market of 0.501 indicates, by Cohen, that the degree of association is large. The result can also be seen visually from Figure 5.13. As there is no way telling if the CDS spreads indicate the true measure of credit quality

the relatively high rank consistency is still a confirmation that the *5-Factor Model* measure the credit quality similarly to how the market values it. The *5-Factor Model* can therefore be used to find approximate CDS spreads for companies for which there aren't any CDS contracts outstanding and also, potentially, in order to identify mispriced CDS contracts.

Although the *5-Factor Model* has similar ranking as the US CDS spreads the same can unfortunately not be said for the Nordic CDS spreads. One plausible explanation is that the US CDS market is more liquid. The lower sample size for the Nordic market should also be mentioned as a limiting factor for the model's performance.

6.3 Model Building

6.3.1 Significance vs Intuition

As may have been noted by now the model building method prioritises economical intuition over perfect data fit. The goal of the model building section is to construct a model which makes sense for the practitioner and which hopefully captures the true underlying nature of credit events.

Having non-intuitive measures will make the model prone to over-fitting the sample. For example, significance of variables such as size or age could be due to identification of potential sampling bias. Although, there is an intuition behind larger companies having more optionality in situations of distress. Age, on the other hand, is seen as quite arbitrary, as it is ill-defined whether to take the starting year of the company as the year the company was launched on the stock market or when the company was founded. M&A activities further dilutes the exactness of this already fuzzy measure of distress.

6.3.2 Including Interaction Terms

When conducting the multivariate correlation analysis interaction terms are considered to be included where *high but not very high* correlations are identified. Although such terms could improve the performance of the model, all such terms are nevertheless disregarded for three main reasons. The first being that the economic intuitiveness of the model could be compromised. Secondly, the relationships are not clear, and in extension it is difficult to concisely define them as new covariates. Thirdly, and most importantly, from a practical point of view, complex covariates can be difficult for the end user of the model to appreciate. Such complex covariates could make the end user doubt the validity of the model.

6.3.3 Market and Sentiment Based Variables

Ideally a model should function both for private and public companies. A model which makes no note of whether the company is public or private is applicable to a wider range of companies. MCAP is the only variable in the data-gathering process which is *forward looking*. Market practitioners' sentiment determines the level of the stock price for the company, along with other fluctuating factors such as supply, demand and liquidity. A model which does not include sentiment based variables is more apt to function as an objective sanity check against possibly unsound market movements. It is expected that the soundness of a non-sentiment based approach is especially rewarding in times just before an emerging crisis. Often in situations of crisis, the market sentiment diverges from what is suggested by the underlying data.

6.3.4 Using Ratios with Different Updating Frequencies

A possible issue in ratio calculations concerns different updating frequencies of numerators and denominators. Say for example that the ratio $R = A/B$ where A is updated daily and B is updated yearly. Furthermore, for notational convenience assume that A was updated at time T and that B was updated at time $T - 364$.

Should one then for calculation of $R(T)$ use $A(T)$ and assume that $B(T) = B(T - 364)$, or should one use $A(T - 364)$ coupled with the most recent known value of B , $B(T - 364)$? There is more information for A available, but B is likely to have changed in the meantime. One faces the choice between information neglect in A , by throwing away new information known at T , and forced constant extrapolation of B . No good solution to this intricacy is presented in this report.

6.3.5 Exclusion of Market Capitalization

There are three main issues concerning the use of MCAP. Firstly, it limits the model usage to only be applicable for public companies. Secondly, MCAP is the only variable in the data that has an updating frequency which differs from the other variables, see discussion in Section 6.3.4 for an explanation of this issue. Thirdly, if MCAP is included in a model then the volatility of a company's stock price can influence the probability of default substantially from day-to-day and this feels unnatural.

As a final note we believe that the strength of MCAP to Total Liabilities, seen in this report, and previously (Altman, 1968), is explained to a great deal by the fact that it is often the only forward looking variable included. A model which functions almost as well but without the use of MCAP, is seen as a great indicator of the model's ability to function as a risk evaluation tool, see discussion in Section 6.3.3.

6.3.6 Traditional Stepwise Inclusion/Exclusion

Initially a traditional stepwise inclusion/exclusion method based on Deviance, AIC or BIC was considered. But these methods are based on asymptotic results, only true if the number of observations are much greater than the number of parameters estimated in the model. Which is not obvious to be true for all subsets of the 63 ratios in this report. Furthermore, fitting of better and better models, according to the named criteria, is not obvious to yield models of increasing discriminative power. The standard criteria were not deemed appropriate for the type of optimisation that is considered in this report and therefore a custom stepwise algorithm is built. The constructed algorithm is one of many possible choices. The resulting model is of course expected to vary with the choice of optimisation criteria. The chosen criterion is deemed appropriate, and more importantly the resulting model (the *5-Factor Model*) performs well out-of-sample, which justifies the adequacy of the algorithm and the chosen selection criterion.

6.3.7 Bias Correction

By implementing finite sample bias correction techniques the resulting probabilities are more realistic with respect to the true probabilities of experiencing credit events (seeing many companies close to 100 % probability, as in Ohlson (1980), is not only unlikely but also unreasonable). This means that the finished *5-Factor Model* is not only viable for relative, but also absolute credit quality assessment. From a practical point of view this is a nice feature.

As seen from the results in Section 5.1.7 the impact from rare event bias reduction is in many cases larger than 10 % of the MLE parameter values (in one case even 15.3 %). This is a substantial effect, and reducing this rare event bias is seen as a great addition from both a practical and theoretical point of view.

Calculating the uncertainty in β is of more theoretical interest than it is useful. But an interesting effect in Figure 5.4 is visible as, although the credit event companies have larger average impacts, the largest effects are seen in the non-credit event data. This is unexpected, but could simply be due to the larger sample size, which makes larger deviance more likely to occur, eventually.

6.3.8 Fraudulent Accounting

All credit event (or bankruptcy) prediction models considered in this thesis are homogeneous in that they rely on accurate accounting data for them to be usable. There are multiple sources of uncertainty in accounting data, to name a few (1) The company or other providers of information can by mistake enter incorrect values, but if this error is large it is likely to be identified

and adjusted for by the company itself, the practitioner of the model or in model building approaches possibly by winsorisation, (2) The corporation may purposely enter faulty numbers into their accounting data to improve its perceived financial health^b. The first, more innocent source of error, is unbiased since a positive and negative impact on the probability of default is equally likely. The second, more malicious source of error, is not only tremendously difficult to identify but also biased, since such fraudulent behaviour is more likely to give rise to a lower, rather than higher, probability of experiencing a credit event. Companies that have been registered for conducting fraudulent accounting are excluded from the data set construction in this thesis.

The second type of error is commonly referred to as *Fraudulent Accounting*. As an example consider the Enron case where three types of fraudulent behaviour, or *shenanigans*, are recorded in Schilit and Perler (2010). The three types, along with a non-exhaustive list of what each type includes, follows, (1) *Earnings*: Recording Revenue Too Soon and Boosting Income Using One-Time or Unsustainable Activities; (2) *Cash Flow*: Shifting Financial Cash Inflows to the Operating Section, Shifting Normal Operating Cash Outflows to the Investing Section and Inflating Operating Cash Flow Using Acquisitions or Disposals; and (3) *Key Metrics*. (Schilit & Perler, 2010)

6.4 The Model in Practice

6.4.1 How to Use the Resulting Model

The *5-Factor Model* is suggested to be used together with common sense and complementary qualitative indicators of financial health. The model is intended to be used for three main purposes, (1) For approximate relative credit quality assessment of companies; (2) For approximate absolute credit quality assessment; (3) For a classification of companies into groups with high and low risks of experiencing credit events. Where (1) and (2) jointly or in isolation can be used to sanity check credit rating and CDS data.

Relative Credit Quality Assessment

If an ordinal ranking of companies based on the risk of experiencing credit events is desired then the discriminative function in Equation 6.1 alone fulfils this purpose.^cThe conversion to probabilities maintains the same ranking, and is therefore unnecessary.

^cThe model is in large expected to perform equally well for another population-wide default rate τ but as the only model tested for rank consistency is the one using $\tau = 2.5\%$ this is the model suggested to be used for relative credit quality assessment.

Absolute Credit Quality Assessment

If the actual probabilities of default are desired then the inverse logit function in Equation 6.2 needs to be used to converse the log-odds from Equation 6.1 to probability estimates. It is important to choose the model coefficients corresponding to the anticipated population-wide default rate τ in Table 5.12.

Classification Assessment

If one needs a crude classification, for companies that are likely or unlikely to experience credit events, then it is suggested to compare the probabilities to the appropriate cutoff retrieved from Table 5.12.

7

Suggestions for Further Research

Science never solves a problem without creating ten more.

– George Bernard Shaw

For variables that have non-monotone ratio-to-default probability relationship, it would be interesting to see if the effect of the *Variable Splitting* technique presented in Section 6.1.3 will yield significance for both the positive and the negative part of the *non-monotone ratios*. This could be an interesting alternative to the approach of Discontinuity Correction utilised in Ohlson (1980).

It would also be of interest to study if a certain type companies are systematically misclassified by the *5-Factor Model*, is there for example a black spot for specific industries, sectors or other natural classifiers of companies.

An examination of the *5-Factor Model*'s rank consistency over time in relation to CDS and Rating data would be interesting to investigate. One natural question concerns whether the ratio-coefficients for the *5-Factor Model* are relatively constant over time. Another emerging dilemma concerns whether the model's ranking has predictive power, for upcoming changes in ratings, or is the rank differences mainly explained by noise in the accounting data? For CDS contracts, one could instead consider a comparison of the implied probabilities of default from CDS contracts to the model output probabilities of experiencing credit events.

As a final remark it would be interesting to evaluate a trading strategy based on the output of the model either in isolation or compared to CDS and credit rating data. Such a strategy could be based on large ranking differences.

8

Conclusion

All models are wrong, but some are useful.

– George E.P. Box

Financial institutions have a direct need of accurate credit quality assessment for risk management and for asserting long-term profitability. A rising sense of urgency is furthermore driven by a steady inflow of new regulations. Market practitioners try to make sense of the asymmetric information issue by turning to the many available sources of credit quality information. Two of the more well known sources are rating agencies and the CDS-market. These sources do however both have their own inherent problems and sources of bias and noise. An accurate and objective model can function as a sanity check for when the bias or the noise takes the upper hand. If the same model also functions where ratings and CDS-data are not updated or available, then assessments of credit qualities are obtainable for a larger amount of companies.

Previous researchers have attempted to formulate quantitative, arguably objective, models, and have come a long way in terms of dichotomous classification of bankruptcies. The performance of two of these models, namely Altman's Z-score and Ohlson's O-score, are in this thesis evaluated, and shown to have acceptable accuracy. A re-estimation of their models to a more recent data set show that the predictive performance is increased compared to the original models.

Due to changes in business climate and company fundamentals, the best predictors of credit quality may be completely different today compared to when the previous researches designed their models. This suggests a need of consideration for alternative models that potentially use different predictors. A model constructed in this thesis performs better than previous models, and this is an indication that the best predictors indeed have changed over recent years. Some of the previous models' ratios are however still significant and efficient predictors which indicates the quality of previous research and that the characteristics of companies about to experience credit events are

not altogether different.

Previous models are only able to provide a classification, or ordinal ranking, of companies based on the likelihood of experiencing default. Techniques that reduce sample and rare-event bias are successfully applied in the model building stage and as a result the constructed model in this thesis has more accurate and more realistic probability estimates than earlier models. The constructed model can, therefore, in addition to relative, be used for absolute credit quality assessment.

The constructed model is shown to be rank consistent with rating and CDS data. This suggests that the model can be used for setting proxy ratings or CDS-spreads, which can be used as sanity checks for CDS price movements that seem out-of-control or for controlling ratings that seem faulty.

The standardised, quantitative and objective model resulting from this thesis is applicable to a wide range of companies; private, public, rated and unrated alike. As providers of the first model with realistic and rank-consistent probabilities, we believe to have come close to capturing the true nature of rare credit events.

Appendices

Appendix A

Construction of Non-Credit Event Sample

In order to find the companies for the non-credit event population, a subset of all available companies in the Bloomberg terminal was considered, through the “Equity Search (EQS)” tool. In the EQS tool it is possible to apply filters and narrow down the number of companies. The filters that were used (on 2016-04-06) are:

- Trading Status: Active
- Security Attribute: Show Primary Security of Company Only
- State of Domicile: United States
- Sector (GICS): -Financials
- Exchanges: United States
- \$100 million \leq CY2014 Total Assets \leq \$50,000 million
- Moody’s Ratng has data
- OR S&P Issuer Rating has data
- OR Fitch Rating has data

1,120 companies, active at the time of the search, are found after applying the filters. Companies that by Bloomberg are considered Financials, REIT or Real Estate companies were excluded. The reason to only choose companies that had total assets between \$100 million and \$50 billion in 2014 is due to the asset size of the companies in the credit event sample. S&P and Fitch, in addition to Moody’s, are included to increase the size of the sample. In the next section, the number of companies is reduced by removal of those that have experienced credit events.

In order to find out if any of the companies have experienced credit event a “Fixed Income Search (SRCH)” was performed in Bloomberg. The SRCH tool works similarly to the EQS tool in that you apply filters to

Bloomberg’s database, but now to fixed income securities. The screenshot from the Bloomberg Terminal that show the filters that were applied are found in Figure A.1. The “Is Defaulted” filter “indicates if the debt instrument is in default or the issuing entity is in bankruptcy, or both are applicable.”. The search resulted in 3,395 instruments.

1. Select Universe SRCH Search Bankrupt
1,454,161 securities

1) Asset Classes Corporates
2) Sources All Securities

2. Criteria Ask a Question















	Field	Boundaries	Selected Criteria	Matches	
31)	Security Status	Include	Bonds: All	1,454,161	 
32) And	Country of Domicile	Include	(United States of America)	305,590	 
33) And	Is Defaulted	--	Yes	5,526	 
34) And	Sector/Industry Grou	Exclude	(Banks or Commercial Finance or Consumer Fina	3,395	 
35) And	Moody's Rating	Has Data	-	-	 
36) L Or	S&P Rating	Has Data	-	-	 
37) L Or	Fitch Rating	Has Data	-	3,395	 
	And		99) Fields		

Figure A.1: Screenshot from SRCH tool in Bloomberg Terminal. (Retrieved 2016-04-06)

If the tickers of the defaulted debts’ issuer matched a company from the EQS, that company was removed from the non-credit event set. The SRCH screening resulted in removal of 56 companies from the non-credit event set and thus 1,064 companies remained. Next, the tickers of the credit event sample of 654 companies was matched to the non-credit event sample, and 28 companies were removed. Due to the fact that not all companies in the credit-event sample had a ticker, a manual matching of the actual names was conducted, and 30 companies were removed following this matching. In the remaining sample of 1,006 non-credit events 16 lacked an industry assignment, and four of those were financials and therefore removed. The final non-credit event sample consists of 1,002 companies.

Appendix B

Bloomberg Variables & Definitions

Variable	BLOOMBERG FIELD ID
Accounts Receivable Turnover	ACCT_RCV_TURN
Accounts & Notes Receivable	BS_ACCT_NOTE_RCV
Accounts Payable	BS_ACCT_PAYABLE
Accounts Payable Turnover	ACCOUNTS_PAYABLE_TURNOVER
Capital Expenditure/Financial Expenditure	CAPITAL_EXPEND
Cash and Near Cash	BS_CASH_NEAR_CASH_ITEM
Cash From Operating Activities	CF_CASH_FROM_OPER
Cash, Cash Eq. & STI	C&CE_AND_STI_DETAILED
Current Assets	BS_CUR_ASSET_REPORT
Current Liabilities	BS_CUR_LIAB
EBIT	EBIT
EBITDA	EBITDA
Free Cash Flow	CF_FREE_CASH_FLOW
Goodwill	BS_GOODWILL
Inventory Turnover	INVENT_TURN
Long Term Debt	BS_LT_BORROW
Market Capitalization	CUR_MKT_CAP
Net Debt	NET_DEBT
Net Income	NET_INCOME
Quick Ratio	QUICK_RATIO
Retained Earnings	BS_RETAIN_EARN
Revenue	SALES_REV_TURN
Short Term Debt	BS_ST_BORROW
Tangible Assets	TANGIBLE_ASSETS
Total Assets	BS_TOT_ASSET
Total Debt	SHORT_AND_LONG_TERM_DEBT
Total Equity	TOTAL_EQUITY
Total Interest Expense	TOT_INT_EXP
Total Invested Capital	TOTAL_INVESTED_CAPITAL
Total Liabilities	BS_TOT_LIAB2
Working Capital	WORKING_CAPITAL

Table B.1: Bloomberg Formulae used to retrieve data

Variable	Definition
Cash Conversion Cycle	$\text{Acc. Rec. T.} + \text{Inv. T.} - \text{Acc. Pay. T}$
CHIN (Change in Net Income)	$(\text{NI}(t) + \text{NI}(t-1)) / (\text{abs}(\text{NI}(t)) + \text{abs}(\text{NI}(t-1)))$
Cur. Ass. Q. to Cur. Liab. Q.	$\text{Accts Rec} / \text{Accts Pay} \cdot \text{Cur Liab} / \text{Cur Assets}$
Intangibles	Total Assets - Tangible Assets
INTWO	1 if NI negative past 2 years, 0 otherwise
Market Value Total Assets	Total Liabilities + Market Cap
Quick Assets	Accounts Rec + Cash and Near Cash
Size	$\log(\text{Total Assets} / \text{GDP Price Index})$
Solvency without Goodwill	$(\text{Total Equity} - \text{Goodwill}) / (\text{Total Assets} - \text{Goodwill})$

Table B.2: Definition of Variables

Appendix C

Credit Rating Data & CDS Data

C.1 Credit Rating Data

C.1.1 United States

Companies that are listed on any of the exchanges in the United States are considered. Yet again, Financials, REIT and real-estate companies were excluded. There are several different types of ratings that S&P set and we are considering the rating corresponding to Long-Term Issuer ratings by S&P. The filters are listed below and 1,054 companies were found on 2016-04-20

- Trading Status: Active
- Security Attribute: Show Primary Security of Company Only
- Sector (BICS): -Financials
- Exchanges: United States
- S&P LT Local Currency Issuer Credit Rating has data

C.1.2 Nordic Region

The Nordic data set is retrieved in a similar fashion as the United States data set, with the obvious difference that the companies need to be listed on any of the exchanges in Sweden, Norway, Finland or Denmark. 35 Nordic companies were found on 2016-04-20 with the filters below applied.

- Trading Status: Active
- Security Attribute: Show Primary Security of Company Only
- Sector (BICS): -Financials
- Exchanges: Sweden, Norway, Finland, Denmark

- S&P LT Local Currency Issuer Credit Rating has data

C.2 CDS Data

C.2.1 United States

The “Global CDS” (GCDS) tool is used to search for different companies that are reference entities for CDS contracts. In the GCDS tool the following filters are applied:

- Sources: All Sources
- Regions: United States
- Ratings: All Ratings
- Sectors: All but Financials and Government
- Debt Type: Senior
- Currency: USD
- ISDA Definition: 2014
- Tenor: 5Y
- Pricing Source: CBIN

C.2.2 Nordic Region

The GCDS tool from Bloomberg will again be used to find CDS prices for the Nordic companies. The filters that were applied to find the Nordic CDS:s are the following:

- Sources: All Sources
- Regions: Denmark, Finland, Norway, Sweden
- Ratings: All Ratings
- Sectors: All but Financials and Government
- Debt Type: Senior
- Currency: All Currencies
- ISDA Definition: 2014
- Tenor: 5Y
- Pricing Source: CMAL

Appendix D

Flexible Data Set - Pseudo Code

```
load data (this code is ideally only run once)
 $v_1, v_2, \dots, v_n \leftarrow$  load all ratio-constituents, we call these variables  $v_i$ 
for each company  $c$  do
  for each ratio  $r$  do
    calculate ratio value  $c(r) = v_i/v_j$ 
    if something is wrong with the ratioa then
      mark  $c(r)$  as a bad ratio
    end if
  end for
  if all 63 ratios are marked as bad then
    remove company  $c$  from the full sampleb
  end if
end for
```

```
before each analysis
set  $R \leftarrow$  user defined ratio subsets with ratios  $r_1, \dots, r_n$ 
set  $A \leftarrow$  all companies
for each ratio in  $R$  do
  set  $T \leftarrow$  all companies  $c$  for which  $c(r)$  is not marked as bad
  set  $A \leftarrow$  the intersection of  $A$  and  $T$ 
end for
if applicable, perform winsorisation on ratios in  $R$  based on companies
in  $A$ 
conduct analysis on companies in  $A$ 
```

^aIn MATLAB this corresponds to `if(isnan(c(r)), isinf(c(r)) or isempty(c(r))`

^bAll companies not deleted contain some information and will thus be included in at least one univariate or multivariate model.

Appendix E

Stepwise Inclusion/Exclusion - Pseudo Code

E.1 Stepwise Algorithm

```
modelSize ← 0, the number of parameters in the model  
maxSize ← 7, a model shouldn't be too large  
minImprov ← 0.01, minimum required improvement of a new covariate  
prevModel ← 0, no covariates in the model to start with  
modImprov ← 1, set to initialize algorithm  
while modelSize < maxSize AND minImprov < modImprov do  
  for all covariates r not in the model, and not just added do  
    calculate the discriminative power of prevModel when adding r  
    Set rmax to the covariate with best additional disc. power  
  end for  
  add the covariate rmax to prevModel  
  modImprov ← the improvement from the best covariate  
  if any of the covariates in prevModel has a p-value above 0.1 then  
    remove the covariate with the largest p-value  
    mark removed covariate so it will be excluded in the next iteration  
  end if  
  modelSize ← the number of covariates in prevModel  
end while
```

E.2 Algorithm Application to Successive Families

r_1, r_2, \dots, r_n are set to be all covariates in Covariate Family 1

$model_0$ is set to be the resulting model from stepwise algorithm on Covariate Family 1

for $i = 2 : 4$ **do**

$model_i$ is set to be the resulting model from stepwise algorithm considering inclusion of covariates from Covariate Family i to $model_{i-1}$

end for

Appendix F

Descriptive Statistics

In Table F.1 the descriptive statistics, (mean, standard deviation, min, median and max) are presented. The last column indicates how many companies that are included in the analysis, i.e. how many companies that have the corresponding ratio available.

#	Ratio	Mean		Std.Dev.		Min		Median		Max		# Companies	
		C	NC	C	NC	C	NC	C	NC	C	NC	C	NC
1	CFO/TL	0.01	0.21	0.09	0.26	-0.47	-0.16	0.02	0.15	0.29	1.71	127	239
2	Cash Ratio 1	0.25	0.50	0.42	0.63	0.00	0.00	0.10	0.27	2.67	3.16	127	241
3	Cash Ratio 2	0.25	0.63	0.40	1.02	0.00	0.00	0.10	0.28	2.45	6.85	122	239
4	CASH/Revenue	0.08	0.13	0.12	0.23	0.00	0.00	0.04	0.05	0.67	1.56	127	240
5	CASH/Ta	0.07	0.11	0.07	0.11	0.00	0.00	0.04	0.07	0.30	0.43	127	241
6	FCF/TL	-0.04	0.09	0.13	0.23	-0.61	-0.54	-0.02	0.07	0.26	1.05	127	239
7	Int. Serv. C.	0.59	-2.34	4.17	5.34	-10.99	-28.53	0.58	-1.05	18.68	13.34	125	243
8	EBIT margin	-0.03	0.09	0.24	0.17	-1.28	-0.73	0.02	0.09	0.33	0.66	127	247
9	EBIT/TA	0.01	0.13	0.15	0.17	-0.76	-0.39	0.03	0.11	0.27	0.76	127	240
10	EBIT/TIntExp	-0.02	12.43	1.79	44.24	-9.75	-91.00	0.37	3.71	5.08	335.94	119	225
11	EBITDA margin	0.08	0.16	0.20	0.19	-0.89	-0.60	0.09	0.14	0.61	0.78	127	244
12	EBITDA/ND	0.08	0.54	0.22	2.99	-1.09	-12.15	0.11	0.34	0.72	23.60	127	238
13	EBITDA/TD	0.10	1.93	0.19	9.36	-0.53	-2.51	0.11	0.40	1.46	77.68	127	230
14	EBITDA/TIntExp	0.95	20.26	1.73	68.76	-7.00	-33.18	1.06	6.15	7.20	595.61	119	222
15	NI/Revenue	-0.38	-0.01	0.64	0.24	-3.34	-1.45	-0.13	0.04	0.08	0.43	127	247
16	NI/TA	-0.30	0.02	0.43	0.16	-2.51	-0.67	-0.15	0.04	0.07	0.45	127	240
17	NI/TD	-0.29	0.81	0.59	4.72	-5.82	-5.87	-0.13	0.11	0.13	36.57	127	231
18	NI/TE	-0.32	0.05	4.28	0.57	-26.35	-1.88	0.14	0.09	7.65	3.06	127	240
19	NI/TL	-0.18	0.04	0.20	0.25	-0.86	-0.98	-0.10	0.06	0.11	0.85	127	240
20	RE/TA	-0.90	0.01	1.21	0.58	-7.05	-3.15	-0.47	0.10	0.35	0.95	127	239
21	INT/TE	0.67	0.84	5.10	2.34	-8.13	-1.60	0.00	0.38	33.74	21.79	116	232
22	LTD/TA	0.87	0.45	0.92	0.50	0.00	0.00	0.70	0.34	5.06	3.06	127	240
23	LTD/TD	0.70	0.87	0.42	0.22	0.00	0.00	0.97	0.96	1.00	1.00	127	232
24	LTD/TInvCap	0.90	0.41	0.80	0.42	0.00	0.00	0.80	0.36	3.44	3.46	127	240
25	OENEG	0.73	0.27	0.44	0.44	0.00	0.00	1.00	0.00	1.00	1.00	127	241
26	STD/TD	0.30	0.13	0.42	0.22	0.00	0.00	0.03	0.04	1.00	1.00	127	230
27	STD/TInvCap	0.40	0.05	0.74	0.09	-0.01	0.00	0.03	0.01	4.01	0.55	127	238
28	Solvency	-0.25	0.60	0.74	0.52	-3.69	-0.62	-0.11	0.51	1.21	3.07	127	241
29	Solv. w/o G	-0.39	0.25	0.61	0.36	-2.41	-1.33	-0.25	0.31	0.59	0.89	113	226
30	TE/LTD	-31.19	14.06	192.31	64.54	-1308.55	-0.75	-0.12	1.50	384.95	506.08	111	228
31	TE/ND	0.11	0.62	1.06	12.71	-0.88	-95.81	-0.11	1.03	6.79	46.21	127	240
32	TE/STD	-16.07	256.63	180.35	856.31	-991.20	-11.63	-0.23	28.72	652.88	6665.79	112	187
33	TE/TD	-0.04	7.55	0.70	29.46	-4.68	-0.74	-0.11	1.31	3.09	208.17	127	232
34	TE/TL	-0.04	1.04	0.34	1.47	-0.65	-0.35	-0.08	0.69	1.53	10.14	127	241
35	TL/TA	0.37	-0.28	0.52	0.57	-0.79	-2.15	0.29	-0.31	1.92	1.42	127	241
36	Size	6.51	6.82	1.09	0.89	4.44	3.95	6.37	6.85	8.99	8.71	127	241
37	CA/Revenue	-1.16	-0.96	0.61	0.73	-2.66	-2.89	-1.17	-1.04	0.80	1.73	127	240
38	CA/TA	-1.07	-0.85	0.71	0.75	-2.94	-2.96	-1.02	-0.53	-0.03	-0.03	127	241
39	CL/CA	0.10	-0.56	1.07	0.58	-1.62	-2.14	-0.24	-0.62	2.68	0.85	127	241
40	CL/TA	-0.97	-1.41	1.01	0.69	-3.28	-3.48	-1.03	-1.39	1.55	-0.12	127	241
41	QR	-0.81	-0.08	1.10	0.79	-3.51	-2.32	-0.49	0.00	1.13	2.07	127	241
42	WC/Revenue	-0.30	0.23	1.03	0.41	-4.53	-0.34	0.06	0.16	1.36	2.71	127	240
43	WC/TA	-0.21	0.22	0.80	0.22	-4.25	-0.29	0.06	0.20	0.65	0.69	127	241
44	APT	11.93	13.04	8.84	15.05	1.45	0.03	9.75	9.46	60.63	104.22	121	214
45	AR/AP	2.02	2.33	2.39	2.69	0.00	0.00	1.48	1.72	15.48	17.19	126	232
46	AR/Revenue	0.11	0.15	0.07	0.19	0.00	0.00	0.11	0.13	0.37	1.64	127	239
47	ACT	2.29	2.28	0.86	0.92	0.95	-0.33	2.06	2.09	5.95	5.24	118	232
48	Cash C. C.	35.05	27.38	86.80	65.71	-45.28	-47.66	8.14	7.38	557.56	450.59	92	182
49	IT	2.28	2.06	1.16	1.17	-0.16	-0.64	2.17	1.84	5.74	5.73	99	190
50	IT/WC	-0.19	-0.05	2.13	1.69	-15.29	-9.17	0.02	0.02	5.61	8.68	99	190
51	QA/CL	-0.85	-0.15	1.09	0.77	-3.51	-2.33	-0.54	-0.07	1.09	1.46	127	240
52	QA/Revenue	-1.91	-1.68	0.84	0.95	-5.49	-4.32	-1.81	-1.62	-0.29	1.03	127	239
53	QA/TA	-1.82	-1.56	0.83	0.84	-4.80	-3.77	-1.65	-1.39	-0.40	-0.25	127	240
54	Revenue/TA	1.41	1.49	0.89	1.05	0.17	0.12	1.31	1.29	4.65	6.00	127	240
55	Revenue/TD	0.09	1.23	0.96	1.34	-2.02	-1.54	0.13	1.08	3.31	5.66	127	231
56	MCAP/TL	-3.05	0.23	1.55	1.07	-6.69	-2.60	-2.87	0.22	0.19	3.21	73	226
57	NI/MTA	-0.21	-0.01	0.21	0.11	-0.83	-0.57	-0.13	0.03	0.02	0.14	73	225
58	TL/MTA	0.91	0.45	0.11	0.21	0.46	0.04	0.95	0.45	1.00	0.93	73	226
59	WC/MCAP	-27.47	0.34	118.54	0.54	-890.93	-0.32	0.03	0.21	101.83	3.36	73	226
60	CHIN	-0.36	-0.05	0.60	0.59	-1.00	-1.00	-0.43	0.01	1.00	1.00	125	243
61	CAQ/CLQ	4.37	1.43	11.08	1.87	0.00	0.00	1.32	0.90	77.16	10.80	126	232
62	INTWO	0.68	0.14	0.47	0.34	0.00	0.00	1.00	0.00	1.00	1.00	125	243
63	Net S. C.	0.04	0.17	0.29	0.36	-0.43	-0.53	-0.01	0.10	1.57	2.12	125	243

Table F.1: Descriptive Statistics for the ratios. Based on the Credit Event Companies (C) and Non-Credit Event Companies (NC) in the estimation set that have the ratios available

Appendix G

Covariate Family Assignment

Subset	#	Ratio Short Name	Welch	K-S	Univariate log	
Covariate Family 1	1	CFO/TL	R	R	R	
	2	Cash Ratio 1	R	R	R	
	3	Cash Ratio 2	R	R	R	
	5	Cash/TA	R	R	R	
	6	FCF/TL	R	R	R	
	8	EBIT margin	R	R	R	
	9	EBIT/TA	R	R	R	
	10	EBIT/ThntExp	R	R	R	
	11	EBITDA margin	R	R	R	
	12	EBITDA/Net Debt	R	R	R	
	13	EBITDA/TD	R	R	R	
	14	EBITDA/ThntExp	R	R	R	
	15	NI margin	R	R	R	
	16	NI/TA	R	R	R	
	17	NI/TD	R	R	R	
	19	NI/TL	R	R	R	
	20	RE/TA	R	R	R	
	28	Solvency	R	R	R	
	29	SwG	R	R	R	
	30	TE/LTD	R	R	R	
	32	TE/STD	R	R	R	
	33	TE/TD	R	R	R	
	34	TE/TL	R	R	R	
	37	log (CA/Rev)	R	R	R	
	38	log (CA/TA)	R	R	R	
	41	log (QR)	R	R	R	
	42	WC/Rev	R	R	R	
	43	WC/TA	R	R	R	
	51	log (QA/CL)	R	R	R	
	52	log (QA/Rev)	R	R	R	
	53	log (QA/TA)	R	R	R	
	55	log (Rev/TD)	R	R	R	
	56	log (MCAP/TL)	R	R	R	
	57	NI/MTA	R	R	R	
	59	WC/MCAP	R	R	R	
	60	CHIN	R	R	R	
	63	Net Sales Change	R	R	R	
	Covariate Family 2	4	Cash/Rev	R	-	R
		7	IntSerCR	-	R	R
		22	LTD/TA	-	R	R
		23	LTD/TD	R	-	R
		24	LTD/TotIntCap	-	R	R
		25	OENEG	-	R	R
		26	STD/TD	-	R	R
		27	STD/TotIntCap	-	R	R
		35	log (TL/TA)	-	R	R
		39	log (CL/CA)	-	R	R
		40	log (CL/TA)	-	R	R
		46	AR/Rev	R	-	R
		58	TL/MTA	-	R	R
		61	CAQ/CLQ	-	R	R
	62	INTWO	-	R	R	
	Cov. Family 3	18	NI/TE	-	R	-
		21	INT/TE	-	R	-
		31	TE/ND	-	R	-
		36	Size	-	-	R
		49	log(TT)	-	-	R
		50	IT/WC	-	R	-
	C. Fam. 4	44	APT	-	-	-
		45	AR/AP	-	-	-
		47	log (ART)	-	-	-
		48	Cash Convession Cycle	-	-	-
		54	Rev/TA	-	-	-

Table G.1: Initial Covariate Families, R means that the corresponding test's H_0 is rejected for that ratio

Appendix H

Correlation Matrix

In the following page the correlation matrix for the 63 ratios is presented. Red markings indicates *very high* correlation and yellow markings indicate *high but not very high* correlation.^a

^aSpecial thanks to Magnus Wiktorsson for the implementation in L^AT_EX.

Appendix I

Representative Visual Examples

I.1 Orders of Magnitude

Figure I.1 is a representative example of where the visual analysis does not yield any useful information. One possible explanation is that there is a difference in order of magnitude within both ratios. This would suggest taking the logarithm of the ratios but as EBIT and EBITDA both go negative frequently this is not possible. Of course one could consider other transformations but this would likely yield something less intuitive and possibly more difficult to interpret. Most visualization plots have this order of magnitude issue and the conclusion in all of them are that no intuitive and interesting patterns are found. Furthermore, no ratios are removed due to order of magnitude.

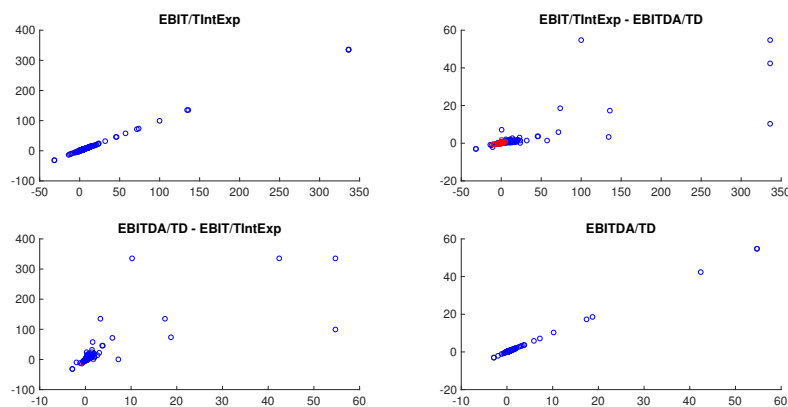


Figure I.1: Plot of EBITDA/TD and EBIT/TIntExp, red indicate credit event & blue indicate non-credit event

I.2 Self-Explanatory Regions

Some situations look like Figure I.2. Both of the ratios appear to go negative for the credit event companies, which could be formulated as an interaction term. But at a closer look, in almost all cases, the appearance has a natural explanation. The ratios share numerator but have distinct denominators, as Net Income turns negative it does so for both the ratios, and as both denominators are strictly positive but not identical there is some scattering effect in the top right and the bottom left plots. An indicator function for these ratios would thus only capture the Net Income effect, therefore the idea was disregarded. Furthermore, as no conclusion of which of the two ratios to include is made, solely based on the correlation and visual analysis, both ratios are kept for further analysis.

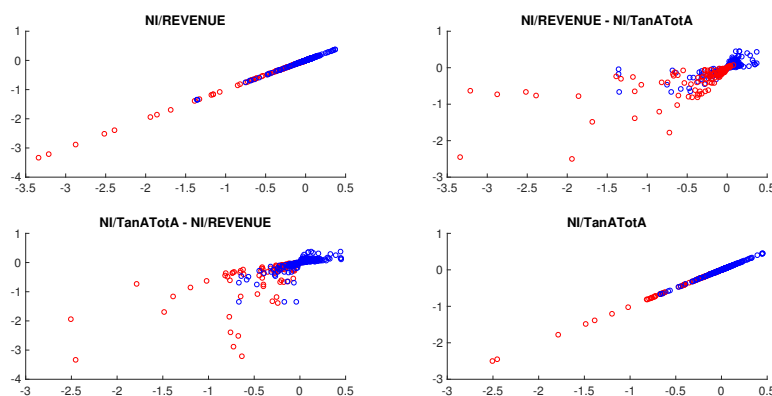


Figure I.2: Illustrating self-explanatory region

I.3 Non-evident Patterns

Figure I.3 is an interesting plot, albeit not very informative. The two ratios are clearly interlinked but the behaviour of credit event companies versus non-credit event companies is quite noisy in the scatter plots. Cash to Current Liabilities obviously stands in close relation to the Quick Ratio, as it is one of its components. The non-linear appearance of the scatter-plots is explained by the other components of the Quick Ratio, i.e. Short Term Investments and Accounts Receivables. We are unable to formulate any relation, but there is an apparent relationship between the two ratios. Nor were we able to remove any of these ratios based on visual analysis. ^a

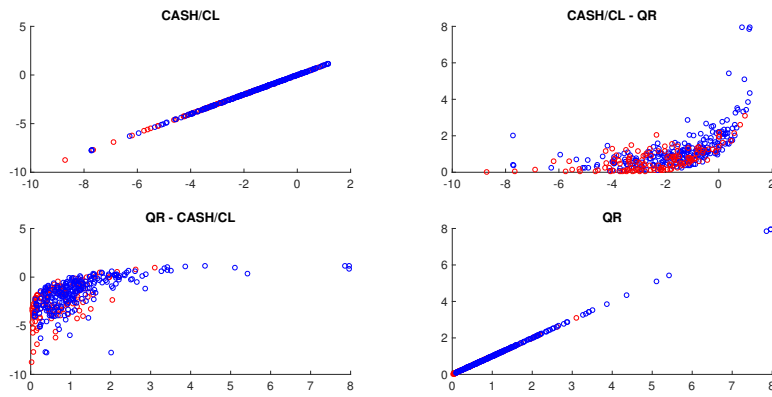


Figure I.3: Illustrating non-evident patterns

^aThis ratio-pair did however happen to have correlation above 0.8 in absolute value. Therefore, since both ratios belong to Covariate Family 1, Cash to Current Liabilities was removed based on economic intuition (the concept Quick Ratio belong to standard financial vernacular).

Appendix J

Histogram for Final Ratios

Figure J.1 contains a histogram of the ratios in the Final Ratio model. In the EBITDA/TIntExp plot most credit events are centered around zero, and thus barely visible. A green background indicates that the ratios have significantly separate means according to Welch's t-test.

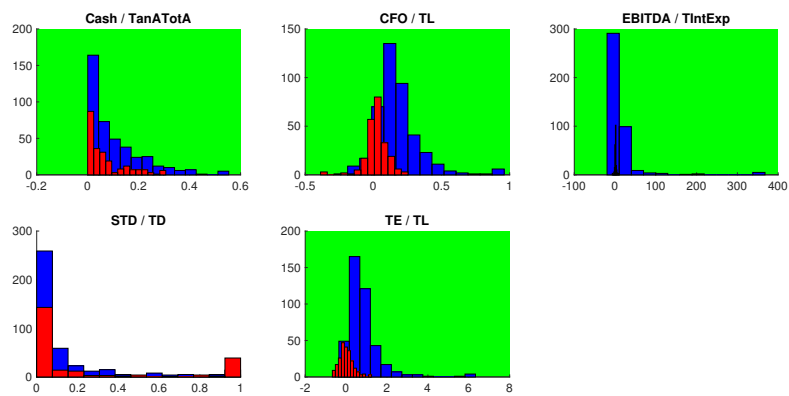


Figure J.1: Histograms of the ratios in the final ratio-model

Appendix K

Re-Estimated Altman & Ohlson

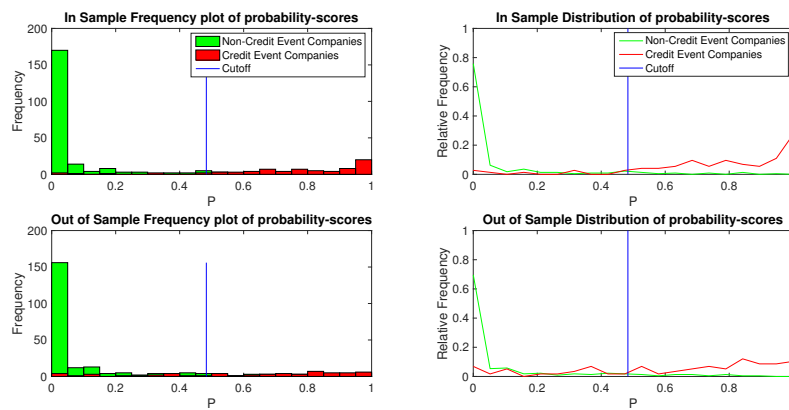


Figure K.1: In- and out-of-sample frequency and distribution plots for the probabilities obtained from the re-estimation of Altman's model

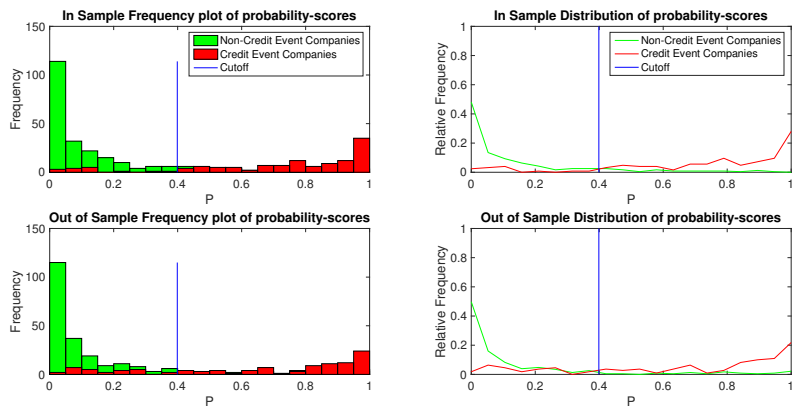


Figure K.2: In- and out-of-sample frequency and distribution plots for the probabilities obtained from the re-estimation of Ohlson's model

Bibliography

- Agresti, A. (2007). *Introduction to categorical data analysis* (2nd ed). Wiley Series in Probability and Statistics, V.423. Wiley.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. Retrieved from <http://www.jstor.org/stable/2978933>
- Altman, E. I. (2000). Predicting financial distress of companies: revisiting the z-score and zeta® models. Retrieved from <http://pages.stern.nyu.edu/~ealtman/Zscores.pdf>
- Amemiya, T. & Vuong, Q. H. (1987). A comparison of two consistent estimators in the choice-based sampling qualitative response model. *Econometrica*, 55(3), 699–702. Retrieved from <http://www.jstor.org/stable/1913609>
- Arakelyan, A. & Serrano, P. (2012). Liquidity in credit default swap markets. Retrieved from http://dee.uib.es/digitalAssets/194/194560_arakelyan.pdf
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71–111. Retrieved from <http://www.jstor.org/stable/2490171>
- Berk, J. & DeMarzo, P. (2013). *Corporate finance* (Global Edition). Pearson Higher Education.
- Breiman, L. (1968). *Probability*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics. Retrieved from <https://books.google.se/books?id=ylyyoUQXkeAC>
- Brown, M. B. & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346), 364–367. Retrieved from <http://www.jstor.org/stable/2285659>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed). Wiley Series in Probability and Statistics. Routledge.
- Dobson, A. J. & Barnett, A. (2007). *An introduction to generalized linear models* (3rd ed). Wiley Series in Probability and Statistics, V.423. Chapman & Hall.
- Faello, J. (2015). Understanding the limitations of financial ratios. *Academy of Accounting & Financial Studies Journal*, 19(3), 75–85. Retrieved

- from <http://ludwig.lub.lu.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=113046436&site=eds-live&scope=site>
- Freedman, D. A. (2006). On the so-called "huber sandwich estimator" and "robust standard errors". *The American Statistician*, 60(4), 299–302. Retrieved from <http://www.jstor.org/stable/27643806>
- Geyer, C. J. (2003). Lecture notes to statistics 5102. University of Minnesota, Twin Cities. Retrieved from <http://www.stat.umn.edu/geyer/old03/5102/notes/fish.pdf>
- Givens, G. H. & Hoeting, J. A. (2013). *Computational statistics* (2nd ed). Wiley Series in Computational Statistics. John Wiley & Sons, Ltd.
- Harrell, F. (2015). *Regression modeling strategies* (2nd ed). Springer Series in Statistic. Springer International Publishing.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70. Retrieved from <http://www.jstor.org/stable/4615733>
- Hosmer, D. W., Lemeshow, S., & Sturvidant, R. X. (2013). *Applied logistic regression* (3rd ed). John Wiley & Sons, inc.
- Kendall, M. G. & Smith, B. B. (1939). The problem of m rankings. *The Annals of Mathematical Statistics*, 10(3), 275–287. Retrieved from <http://www.jstor.org/stable/2235668>
- King, G. & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163. Retrieved from <http://www.jstor.org/stable/25791637>
- Kleinberg, J. & Tardos, E. (2005). *Algorithm design*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- Levene, H. (1960). Robust tests for equality of variance. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, & H. B. MannI (Eds.), *Contributions to probability and statistics: essays in honor of harold hotelling* (pp. 278–292). Stanford: Stanford University Press.
- Löffler, G. & Posch, P. N. (2007). *Credit risk modelling using excel and vba* (2nd ed). Wiley Finance. John Wiley & Sons, Ltd.
- Manski, C. F. & Lerman, S. R. (1977). The estimation of choice probabilities from choice based samples. *Econometrica*, 45(8), 1977–1988. Retrieved from <http://www.jstor.org/stable/1914121>
- McCullagh, P. & Nelder, J. A. (1989). *Generalized linear models* (2nd ed). Monographs on Statistics and Applied Probability 37. Chapman & Hall/CRC.
- Moody's Investor Services. (2016, February). *Rating symbols and definitions*. Moody's Publications.
- NIST. (2012). Levene test for equality of variances. Retrieved from <http://www.itl.nist.gov/div898/handbook/>

- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. Retrieved from <http://www.jstor.org/stable/2490395>
- The Concise Encyclopedia of Statistics. (2008). (Chap. Maximum Likelihood, pp. 334–336). New York, NY: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-0-387-32833-1_248
- The Concise Encyclopedia of Statistics. (2008). (Chap. Student Test, pp. 528–532). New York, NY: Springer New York. Retrieved from http://dx.doi.org/10.1007/978-0-387-32833-1_395
- Panchenko, D. (2006). Lecture notes in statistics for applications. Massachusetts Institute of Technology. Retrieved from <http://ocw.mit.edu/courses/mathematics/18-443-statistics-for-applications-fall-2006/lecture-notes/lecture14.pdf>
- Rodríguez, G. (2001). Lecture notes to generalized linear models. Princeton University. Retrieved from <http://data.princeton.edu/wws509/notes/a1.pdf>
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to student's t-test and the mann-whitney u test. *Behavioral Ecology*, 17(4), 688–690. Retrieved from <http://beheco.oxfordjournals.org/content/17/4/688.short>
- Schilit, H. & Perler, J. (2010). *Financial shenanigans: how to detect accounting gimmicks & fraud in financial reports* (3rd ed). McGraw-Hill Education.
- Scott, A. J. & Wild, C. J. (1986). Fitting logistic models under case-control or choice based sampling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(2), 170–182. Retrieved from <http://www.jstor.org/stable/2345712>
- Services, M. I. (2015, March). *Annual default study: corporate default and recovery rates, 1920-2014*. Report Number: 179348. Moody's Publications.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: a simple hazard model. *The Journal of Business*, 74(1), 101–124. Retrieved from <http://www.jstor.org/stable/10.1086/209665>
- Vassalou, M. & Yuhang, X. (2004). Default risk in equity returns. *The Journal of Finance*, 59(2), 831–868. Retrieved from <http://www.jstor.org/stable/3694915>
- Welch, B. L. (1947). The generalization of 'student's' problem when several different population variances are involved. *Biometrika*, 34(1-2), 28–35. Retrieved from <http://biomet.oxfordjournals.org/content/34/1-2/28.short>
- White, L. J. (2010). Credit-rating agencies and the financial crisis: less regulation of cras is a better response. *Journal of International Banking Law and Regulation (UK)*, 25, 170–179. Retrieved from http://www.stern.nyu.edu/sites/default/files/assets/documents/con_039549.pdf

- Xie, Y. & Manski, C. F. (1989). The logit model and response-based samples. *Sociological Methods & Research*, 17(3), 283–302. Retrieved from <http://smr.sagepub.com/content/17/3/283.abstract>
- Zamar, R. (2014). Lecture notes in statistical inference i. Department of Statistics University of British Columbia. Retrieved from <http://www.stat.ubc.ca/~ruben/Stat560Website/Lectures.2015/Lecture10.2014.pdf>
- Zhou, X.-H., Obuchowski, N. A., & McClish, D. K. (2011). *Statistical methods in diagnostic medicine* (2nd ed). Wiley Series in Probability and Statistics. John Wiley & Sons.