LUND UNIVERSITY
School of Economics and Management

Altman versus Merton

Are corporate credit rating changes new information?

by

Anton Andersson

Joakim Brunnhage

June 2016

Master's Programme in Corporate and Financial Management

Supervisor: Håkan Jankensgård

Abstract

With the relative performance between accounting-based and option-based approaches for default prediction being a key subject in previous research, combined with criticism against the rating agencies' timeliness in assigning credit ratings, surprisingly few attempts have been made to investigate the models' usefulness in predicting corporate credit ratings. This thesis is the first to investigate the Merton and Altman's Z-score models' relative performance in predicting future credit rating changes. In addition, any asymmetry in the predictive power for downgrades versus upgrades is examined. The thesis is performed using data from 1,450 non-financial firms rated by Standard and Poor's between 2002 and 2013. Using logit regressions, with rating changes as dependent variables and distances-to-default and Z-scores as independent variables, we find both models to have some predictive power for rating changes within one year but the goodness of fit is mediocre and the marginal effects are low. Although the Z-score shows slightly better results, in terms of percentage correctly predicted outcomes, it is concluded that no clear difference in the relative performance can be found. Meanwhile, it is concluded that the Merton model has more predictive power for downgrades than for upgrades while no such asymmetry can be found for the Z-score. Our results supports the notion of lagged credit ratings, which could be detrimental for the economy at large, and may work as a starting point for building more accurate prediction models to lessen the effect of rating announcements. It is further implied that credit rating agencies could be slower to assign lowered credit ratings, as compared to rating upgrades, which could possibly be explained by the interdependency between the rating agencies and the issuers.

Acknowledgements

We would like to thank our supervisor Håkan Jankensgård at the Department of Business Administration for his helpful advising throughout the writing process, especially in times of trouble and uncertainty. Further, we would like to express our gratitude to Anamaria Cociorva for her guidance when retrieving the credit rating data needed to perform this study.

Table of Contents

1 Introduction

This section introduces the area of interest, starting with a discussion regarding the importance of credit ratings in today's capital markets. Closely followed is a walkthrough of the research background concerning other methods of credit risk before the subject is problematized. Thereafter follows a presentation of the purpose of the thesis, research limitations and a description of the thesis outline.

1.1 Background

Credit ratings have become almost indispensable in today's capital markets as virtually all market participants, including issuers, investors and regulators, rely heavily upon the categorical letter scores assigned by the credit rating agencies (CRAs). Issuers are affected mainly through changes in financing costs (Langohr & Langohr, 2009), and the credit ratings have shown to be the second most important determinant for CFOs when deciding whether or not to issue additional debt (Graham & Harvey, 2001). For investors, credit ratings serve as crucial components in the evaluation process of investment opportunities and finally, to the regulators, credit ratings are essential for determining allowable investment alternatives, for institutional investors, and required capital for banks (Langohr & Langohr, 2009). Credit ratings, based on both quantitative and qualitative aspects, facilitate easier distinction between cherries and lemons and the credit scoring system comes with several advantages. For example, it enables comparison and benchmarking of entities, from different countries and industries, and the monitoring of the development of credit risk over time. As each score corresponds to a certain risk level it also gives an indication of the risk premium one might expect to receive (Ganguin & Bilardello, 2005). In addition, credit ratings have shown to be accurate in predicting corporate default (Orth, 2012).

In essence, the CRAs specialize in mitigating problems relating to information asymmetry between issuers and creditors, a gap that can never fully be bridged. As long as credit ratings carry unique information that is not already embedded in the market prices the CRAs play a pivotal role by adding valuable information to the market, earning its participants' continued attention (Langohr & Langohr, 2009).

1.2 Research Background

During the second half of the 20th century, alternative and more quantitatively oriented methods, for assessing credit risk, have been developed. These can more broadly be divided into accounting-based and option-based models. The accounting-based models rely solely on accounting variables for determining the probability of default while option-based models are based on the option-pricing theory developed by Black and Scholes in 1973 (Tanthanongsakkun & Treepongkaruna, 2008).

The research regarding default prediction using accounting-based models stems from the pioneering work of William Beaver, in 1966, in which he strived to assess the usefulness of a set of financial ratios in predicting corporate failure. Beaver (1966) was closely followed by Altman (1968) whose findings laid the foundation for the famous Altman Z-score model, rigorously adopted in subsequent research, such as Ohlson (1980) and Zmijewski (1984). Almost a decade later, another approach for tackling default prediction emerged with the work of Merton (1974). This model, being option-based, had a forward-looking approach to assessing credit risk rather than the backward-looking groundwork of Beaver (1966). Several researchers have also tried to test the performance of the Merton model, some ending up with alternative models, such as Du and Suo (2007), Bharath and Shumway (2008) and Afik, Arad and Galil (2015).

With both approaches being popularly used for default prediction, resolving the ambiguity in the relative performance of the models has been a key subject for subsequent research. The majority of studies, where among others Hillegeist et al. (2004), Tanthanongsakkun and Treepongkaruna (2008) and Gharghori, Chan and Faff (2006), have shown relatively conclusive results in favor of the option-based approaches.

Seemingly few attempts, however, have been made to use the aforementioned models for explaining actual corporate credit ratings. Of the scarce research attention that exists in this area, most studies have adopted only accounting-based models, Pogue and Soldofsky (1969) and Pinches and Mingo (1973), with only a few, such as Tanthanongsakkun and Treepongkaruna (2008), examining option-based models in this aspect.

1.3 Problem Discussion

If the ratings put forth by the CRAs and the probability of default, as calculated with the aforementioned models, essentially is a measure of the same thing, credit default risk, it seems relevant that there should be a correlation between the measures. Meanwhile, since the beginning of the 21st century, the CRAs have been heavily criticized for failing in its timeliness in assessing credit rating changes. The vast majority of market participants, whom rely on credit ratings, does not fully believe that the CRAs' rating changes timely reflect the subject's changed credit risk (Association for Financial Professionals, 2002). This was especially a problem for the high-profile failures of Enron and WorldCom where both Moody's and Standard and Poor's (S&P) had Enron and WorldCom investment-grade rated until six days and two months respectively before the companies declared bankruptcy, implying that credit ratings are lagged (Langohr & Langohr, 2009). This potential lag could be a result of the credit ratings being both long-term looking and non-procyclical (Amato & Furfine, 2004). The lag could also be due the fact that firms, rated by the CRAs, also constitute the majority of the CRAs revenue base, which causes a potential conflict of interest where the CRAs might think twice before a downgrading. A study by Feinberg, Shelor and Jiang (2004) shows that the interdependency between traditional CRAs, including Moody's and S&P, and issuers cause the CRAs to react slower to changes as compared to independent CRAs.

If credit ratings are lagged, and the measures are correlated, explaining current credit ratings using either accounting-based or option-based models would be futile. Instead, one would have a greater success in using the aforementioned models for predicting *future* credit ratings. Under this assumption, one could thus potentially predict future credit rating *changes*. But why would such a discovery be of interest?

Empirical research has shown that credit rating announcements carry unique information not already embedded in the market prices. Hand, Holthausen and Leftwich showed in 1992 that bond- and stock prices in the U.S. market were affected by announcements of both actual credit rating changes and additions to the Credit Watch list. They also found that rating downgrades showed more reliable results than upgrades. Barely ten years later, Dichev and Piotroski (2001) supported this by showing that rating downgrades also had negative impact on long-run stock returns, while the announcement of rating upgrades had no significant impact. Similar results were found by Freitas and Minardi (2013) for Latin American firms. Brogaard, Koski and Siegel (2015) adopted a different approach and studied the effect of rating changes on the trading volume, a proxy for information content, instead of stock prices. They found that abnormal trading volumes occurred for both rating up- and downgrades, which indicates that rating changes announcements are informative for investors.

Thus, the ability to predict such changes of credit ratings would indicate that one could act before the market, before prices are adjusted to embed the updated credit rating. On the other hand, if widely applied, it could potentially question the value of rating announcements, essentially rendering them obsolete.

As it turns out, only Westerlund and Rebeggiani (2012) have investigated this opportunity prior to this study, however, using only the option-based approach for a sample of 100 firms from the S&P 500 index. However, the area of credit rating *changes* differentiates itself from previous research. Thus, incorporating only option-based models in explaining these changes does not give a complete picture and does not resolve the ambiguity in the relative performance of the models. Because of this, previous results in favor of the option-based approaches does not necessarily still apply. We extend upon the limited knowledge in this research area by re-introducing the so called horse-race between the models.

1.4 Research Purpose

This thesis re-introduces the horse-race between the accounting-based and option-based approaches for default probability, namely Altman's Z-score and the Merton model, for the purpose of investigating their relative performance in predicting future corporate credit rating changes.

In line with previous findings regarding the information content of rating announcements, we further aim to test if any asymmetry, in the abovementioned models' abilities in predicting rating upgrades versus rating downgrades, exists.

1.5 Research Limitations

The ideal timeframe, for fulfilling the purpose, would be to investigate the period of 2002 until today, as explained in section 4.2. Alas, very recent data was unobtainable through the available databases and manually gathering the needed information proved to be too time-consuming given the sample size. Due to this predicament, the timeframe for this study was set to 2002-2013 as this allowed for obtaining the required inputs from available databases.

This thesis is restricted to testing the early models, i.e. Altman's Z-score and the Merton model. Although these models have given rise to a plentitude of alternative models, some being more complex and others combining aspects from both approaches, incorporating these in our study, and more importantly choosing which variants to use, would be somewhat arbitrary. The reasoning behind this lies in the differentiation of our purpose as compared to the bulk of previous research regarding default probability. Thus, the ambiguity in the basic models' relative performance remains unresolved in this new application and this thesis instead aims to lay a foundation from which alternative model testing can arise in subsequent research. Furthermore, Afik, Arad & Galil (2015) show that simple applications of the Merton model outperforms more complex and demanding methods.

Due to the limitations of the Z-score (accounting data is not frequently updated) and that a comparison between the models is the main objective, this thesis is delimited to the minimum observation-frequency of one year.

1.6 Outline of the Thesis

Section 2 presents the theoretical framework, including the credit rating process (adopted by S&P), the Merton model, Altman's Z-score and relevant previous research which lead to our hypotheses, presented in section 3. Section 4 explains the methodology adopted, including a presentation of the sample selection, data collection, variable description and the regression model, and is rounded off with a discussion regarding the validity and reliability of the thesis. Section 5 presents and discusses the results from our empirical findings in connection to the theoretical framework and previous research. The final section presents our conclusions and suggestions for further research.

2 Literature Review

This section includes a description of the theoretical foundation underlying the thesis, including a description of the credit rating process of Standard and Poor's, the Merton model and Altman's Z-score. Further included is a review of previous research in the area, which is summarized in the end of the section.

2.1 Standard and Poor's and Credit Ratings

As the credit rating data obtained consists of credit ratings from S&P, see section 4.3.1, the following segment will present the credit rating process of S&P.

S&P's credit ratings basically constitute their opinion of the rated entity's credit risk, i.e. its ability and willingness to fulfill its financial obligations in time. S&P assigns credit ratings on both corporations and governments as well as on debt-, bond- and notes issues. Their rating opinions are the sum of a thorough analysis by analysts who make use of information acquired from both the entity of interest and publicly available sources (Standard and Poor's, 2014a).

The corporate credit ratings of S&P does not only include an assessment of the financial risk profile of the firm but also a thorough analysis of more soft values like the business risk, consisting of country-, industry- and company-specific risk. The business- and financial risk profile create the so called issuer's anchor which is adjusted with some other factors, including the management factor, before ending up with the final credit rating (Standard and Poor's, 2014b)

The process of assigning credit ratings for S&P takes approximately three to four weeks and starts with the signing of an engagement letter by the issuer requesting a credit rating. An analyst team is then assembled whose first task is to review relevant information and meet with the management of the issuer. Thereafter, the analysis process is initiated where information is analyzed and the outcome of this process is a proposal of a rating to the ratings committee. The committee reviews the analysts' proposed rating and votes on it. Prior to publishing a press release with the issuers rating, S&P informs the issuer on the rationale underlying the proposed credit rating. After the rating has been assigned, S&P continues to survey the company in order to maintain an as current rating as possible, by looking out for events and developments that could lead to a potential up- or downgrade of the credit rating (Standard and Poor's, 2016).

2.2 Merton Model

The theoretical model of distance-to-default (DD) used in this thesis was originally developed by Merton (1974), thus henceforth referred to as the Merton model, and is largely taken from Bharath and Shumway (2008). The model offers an option-based, quantitative approach for assessing the credit risk of a firm, or any entity, that has financial obligations. The Merton model calculates the estimated DD, at a future time T, by adding the relative distance, by which asset value currently exceeds debt, to the expected growth of said asset, during the observed timeframe. The result, which is the expected relative distance between asset value and debt, at time T, is then divided by the standard deviation of asset growth to express the DD in standard deviation terms. To translate the DD into a probability of default, the DD is substituted into a cumulative standard normal distribution which gives the probability that a normally distributed variable stays below any given number.

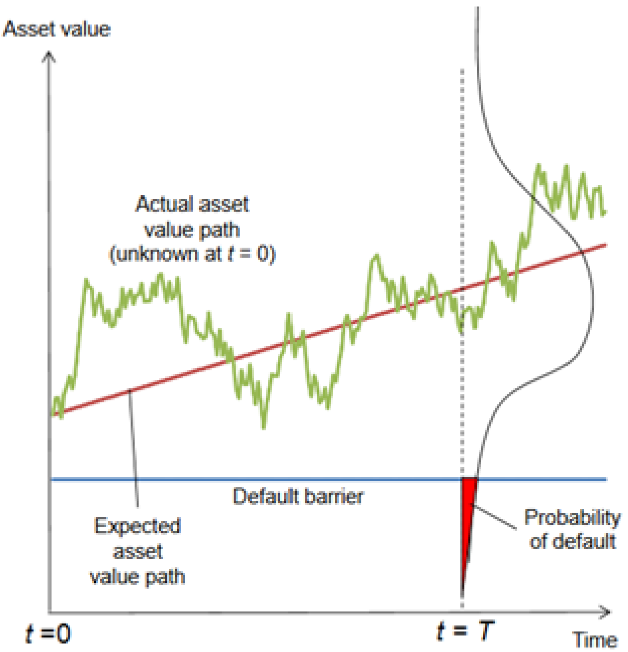

Figure 2.1 – The Merton model graphically explained (Gray & Malone, 2008)

The Merton model exploits the Black and Scholes (1974) option-pricing formula and makes two key assumptions. Firstly, a firm's capital structure consists of only equity and just one non-interest bearing, zero-coupon debt with a maturity at time T. Secondly, the total asset value of a firm follows a geometric Brownian motion,

$$dA_0 = \mu A_0 dt + \sigma_A A_0 dW \tag{2.1}$$

with A_0 being the asset value of the firm, μ being the expected continuously compounded return on assets and σ_A being the asset volatility. Finally, dW is a standard Wiener process. Under these assumptions, a firm's equity can be viewed as a European call option on the firm's assets. This call option analogy is based on the conception that when a firm has met all its financial obligations, the equity holders are the residual claimants on its assets. By setting the strike price of the call option equal to the book value of the firm's liabilities, the call option will only have a value as long as the market value of the firm's assets exceeds the strike price. Thus, the value of equity can be calculated as,

$$E = A_0 N(d_1) - e^{-rT} B N(d_2) \tag{2.2}$$

where E is the market value of equity, B is the face value of the firm's debt, r is the risk-free interest rate and N is the cumulative standard normal distribution function. Here, d_1 and d_2 are given by,

$$d_1 = \frac{\ln\left(\frac{A_0}{B}\right) + \left(r + \frac{\sigma_A^2}{2}\right)T}{\sigma_A \sqrt{T}} \tag{2.3}$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \tag{2.4}$$

The market value of assets and asset volatility is unobservable in reality. However, under the assumptions of the Merton model, with equity being a function of asset value and time, when applying Ito's lemma, it follows that,

$$E\sigma_E = A_0 \sigma_A N(d_1) \tag{2.5}$$

where σ_V is the instantaneous volatility of equity at T=0. The Merton model then uses equation 2.2 and 2.5 to numerically obtain asset value and asset volatility through equity volatility and equity. Thereafter the DD can be calculated as,

$$DD = \frac{\ln\left(\frac{A_0}{B}\right)+\left(\mu-\frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}} \tag{2.6}$$

where μ is an estimate of the expected annual asset growth. The probability of default is then calculated by,

$$PD = N(-DD) \tag{2.7}$$

2.3 Altman's Z-score

The Altman Z-score was introduced in 1968 by Ed Altman in his article "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy". Therein, Altman aimed to assess the ability of a set of financial ratios to predict corporate bankruptcy. Out of 22 potentially useful financial ratios, five were chosen and included in his model firstly based on their popularity in previous research and relevancy to the study. Secondly, he evaluated the relative contribution and correlation of the independent variables and predictive accuracy of the specification, and partly using his own judgement, before ending up with the final specification. This specification included the five following variables combined; Working capital/Total assets (X_1), Retained earnings/Total assets (X_2), EBIT/Total assets (X_3), Market value of equity/Book value of total debt (X_4) and Sales/Total assets (X_5). The function looks as follows:

$$Z - score = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5 \tag{2.8}$$

Using multiple discriminant analysis on a sample of 33 bankrupt and 33 non-bankrupt U.S. manufacturing firms between 1946 and 1965 Altman found the model to be very accurate, predicting over 90 % of corporate bankruptcies. The model showed to be highly reliable when robustness tests where run. To enable predictions of corporate bankruptcy without using complex data procedures Altman took advantage of his findings and put forth thresholds of Z-scores indicating the firms' health status. Firms with a Z-score above 2.99 fell under the "non-bankrupt” category while firms with Z-scores below 1.81 were deemed bankrupt. The area between 2.99 and 1.81 was called the “gray area”.

In the final part of his article Altman presents several additional areas of potential application of the model. He recommends banks to use the Z-score when evaluating loan applicants to lower the costs of monitoring, wherein companies with high scores would be less investigated by the banks and vice versa. For other companies, Altman further suggests using the Z-scores for internal control, to assess the financial health of the firm and predict corporate problems. In addition, the Z-score could also be of use when evaluating investment opportunities.

2.4 Previous Research

2.4.1 Testing Option-Based Models

In their article from 2007, Du and Suo evaluate the performance of structural credit risk models in forecasting default. The study was performed upon a sample of 1,508 U.S. firms between 1989 and 2001 using a dynamic logistic model. The authors found a reduced-form model slightly superior as compared to the original Merton model in its ability to predict default, although with somewhat limited statistical significance. Du and Suo (2007) showed that including equity value into the model greatly improves upon the predictive power.

In Bharath & Shumway’s article, published 2008 in the Review of Financial Studies, the authors examined the accuracy and contribution of the Merton model as a predictor of company defaults. The study was done using data from 1,449 non-financial firms listed on NYSE, AMEX and NASDAQ over the time period of 1980 to 2003. By testing several hypotheses, the authors analyzed the Merton model in several ways. Firstly, looking at hazard models that forecast default and by constructing a reduced form model of the Merton model, it is shown that it is possible to improve upon the Merton model and thus concluded that the Merton model does not provide a sufficient statistic for default prediction. However secondly, by incorporating the default probability calculated by the Merton model’s DD with other covariant variables in hazard models, the authors showed that the Merton model’s default probability is a useful variable for forecasting default due to the functional form of the Merton model. Furthermore, it is concluded that the forecasting ability of the Merton model does not seem very sensitive to the somewhat complex iterative procedure used for calculating the inputs of asset value and asset value volatility. Overall, Bharath and Shumway (2008) argued that the Merton model provides useful guidance for building default forecasting models.

In response to criticism of the Merton approach to evaluate credit risk, from researchers at Moody's, Kealhofer and Kurbat (2001) showed, on a sample of 1,579 North American firms between 1990 and 1999, that a Merton-model-like approach developed by KMV not only outperforms other available default prediction alternatives but also already captures all of the default predictive information in frequently used accounting ratios and Moody's credit ratings. According to Bharath and Shumway (2008), the Merton approach applied here is called the KMV model which allows for various different classes and maturities of the debt, in contrast to the standard Merton model as explained in section 2.2. Furthermore, the KMV model does not use the cumulative normal distribution when calculating the probability of default, but rather Moody's large database to estimate the correct distribution and also needs some adjustments to the accounting information used to calculate the face value of debt. Bharath and Shumway (2008) showed that the Merton DD does a good job in capturing the information in the probability estimates from the KMV model.

Westerlund and Rebeggiani (2012) investigated if the Merton model has any predictive power of changes in Moody's credit ratings, in order to assess whether credit ratings are lagged or not. In addition, the authors tested if there are any differences in the model's ability to predict rating down- and upgrades. The authors used a sample of 100 randomly selected non-financial firms listed on S&P 500 during the time period 2003-2009. Using a logit regression model, and monthly data, the authors only found a 12-month lagged DD measure, out of 14 lagged measures used, to have statistically significant predictive power of credit rating changes. When looking separately at rating downgrades Westerlund and Rebeggiani (2012) found both a 12 and 18-month lag to be statistically significant. No significant results were found when testing for rating upgrades in isolation. The authors concluded that one possible reason for ending up with few statistically significant results is the use of a too small sample of firms with very stable credit ratings, thus ending up with very few rating changes observations.

In their article from 2015, Afik, Arad and Galil evaluated different specifications of the Merton model's abilities for default prediction. The authors used a sample of firms rated by S&P during the time period 1989 - 2012 and also gathered information on defaults between 1990 and 2013, ending up with 306 cases of default for 2,534 firms. The alternative models examined in the study were the iterative estimation (KMV), Bharath and Shumway's (2008) naïve model, the Charitou et al. (2013) model, Down and Out call (DaO), single equation models, the authors' own simple naïve model and the equity volatility estimation model. The authors found that simplified applications of the Merton model outperform the more complex and demanding methods. They further concluded that the DaO is the best performing model compared to the other models, closely followed by the simple naïve model that the authors themselves propose. They ultimately recommended using the later one because of it being more intuitive and simple as compared to the DaO approach.

Cai, Chen and Dong (2015) examined in their article the default prediction ability of three variants of DD measures based upon the Merton model, which differs in either the functional form or the solving methodology. The three approaches examined were the original DD introduced by Merton, the KMV model and the naïve DD proposed by Bharath & Shumway in 2008. The authors compared and analyzed the three approaches' prediction abilities using logit regressions on a sample of 46 defaults for Chinese companies between 2012 and 2014. The authors concluded that the naïve DD is slightly outperforming the other measures, although with somewhat lacking predictability and only 45 % accuracy. To improve performance, the authors suggested adding accounting ratios. For example, they found that EBITDA over total liabilities can supplement the naïve DD measure well.

2.4.2 Testing Accounting-Based Models

Pogue and Soldofsky (1969) tried to assess how well corporate bond ratings can be explained by available financial data on U.S. firms between 1961 and 1966. Of five explanatory variables, the authors found leverage and profitability ratios to have the most significant impact on corporate bond ratings while size, earnings stability and EBIT interest coverage were less influential. The overall explanatory power of the model including all five variables was 56 %, which was more than expected by the authors. Based on their results the authors concluded that a firm's history, rather than its future prospects, to a larger extent can explain major differences in corporate bond ratings.

Pinches and Mingo aimed in their article from 1973 to explain industrial bond ratings using both factor analysis and a multiple discriminant model. The authors used a sample of 180 U.S. firms, 132 forming the original sample and 48 forming the holdout sample, between early

1967 and late 1968. From the factor analysis, Pinches and Mingo (1973) found six variables to be especially important in predicting bond ratings; subordination, years of consecutive dividends, issue size, EBIT interest coverage, long-term debt ratio and return on assets, which were all included in the final multiple discriminant model. The authors found that the abovementioned variables correctly explain almost 70 % of the ratings in the original sample, and around 60 % in the holdout sample.

Gray, Mirkovic and Ragunathan (2006) examined in their article the relationship between Australian credit ratings and a number of different financial ratios and industry-specific variables, expected to have explanatory power of credit ratings. The study was conducted as an ordered probit model where credit ratings from S&P were collected between 1995-2002 for a sample of 361 rating observations, excluding non-investment grade firms, banks and insurance firms. The financial ratios used in the study were interest coverage, cash flow, leverage and profitability. The two industry-specific variables included were systematic risk and industry concentration, proxied by the industry beta and the four-firm concentration ratio respectively, where the last mentioned was measured by how large a part of total industry production the four largest firms in the industry accounts for. Gray, Mirkovic and Ragunathan (2006) concluded that interest coverage and leverage ratios have the largest impact on credit ratings while profitability and industry concentration has less pronounced impact on credit ratings. Their model could correctly explain 61.5 % credit ratings.

2.4.3 Horse-Races

Hillegeist et al. (2004) compared the ability of two accounting-based- and one option-based measure in predicting default on a sample of 756 U.S. bankruptcies between 1980 and 2000. The two accounting-based measures used were the Altman Z-score and the Ohlson O-score while the option-based measure used were their own specification of the Black-Scholes-Merton option pricing model. The authors found that the option-based model performs significantly better in predicting bankruptcy as compared to the two accounting-based measures, even after making modifications of the Altman Z-score and Ohlson O-score, including adjustments for updated coefficients, industry adjustments and a decomposition into lagged levels and changes.

Gharghori, Chan & Faff (2006) evaluated, in their article, the default prediction accuracy of three alternative default-risk models over the time period 1995-2003 for 856 Australian firms. Two of the models tested were option-based and derived from Merton's (1974) finding that equity can be viewed as a call option on the firm's assets. In the first model, equity was modelled as a standard call option while it in the second model was modelled as a path-dependent barrier option. Lastly, the authors evaluated an accounting-based model, using accounting ratios similar to Altman's (1968) Z-score. Here, 13 variables were selected based on previous research adopting accounting-based models. The authors found that the two option-based models perform similarly, as a measure of default risk, and that they clearly outperform the accounting-based model. It is further suggested that the Merton model is preferred due to its simpler characteristics.

Tanthanongsakkun and Treepongkaruna (2008) extended upon the research of Gray, Mirkovic and Ragunathan (2006) and Gharghori, Chan and Faff (2006) and adopted an ordered probit model to examine how effectively an option-based, as compared to an accounting-based, model can explain credit ratings. Tanthanongsakkun and Treepongkaruna (2008) studied a sample of 257 S&P credit rating observations for Australian firms between 1992 and 2003, excluding banks and insurance companies. The accounting-based model included interest coverage and leverage ratios, the variables with the most significant impact on credit ratings according to Gray, Mirkovic and Ragunathan (2006). In the option-based model, probability of default from the Merton model (DLI), firm size and book-to-market were used as explanatory variables. DLI was used in accordance with the findings of Gharghori, Chan and Faff (2006), while the inclusion of firm size and book-to-market ratio was based upon the findings of Vassalou and Xing (2004) who found those variables to be significantly related to default risk. The authors concluded that the option-based model has the most explanatory power in explaining credit ratings, where firm size in isolation is the most significant variable. Further, the option-based model beat the accounting-based model in regards to prediction accuracy with 55.25 % versus 47.47 % correctly predicted credit ratings.

2.4.4 Summary of Previous Research

Authors	Area of interest	Variables	Results
		Testing Option-Based Models	
Du & Suo (2007)	Default prediction	Different specifications of the Merton model	A reduced-form model slightly outperforms the Merton model
Bharath & Shumway (2008)	Default prediction	Different specifications of the Merton model	Merton model provides useful guidance for default prediction
Kealhofer & Kurbat (2001)	Default prediction	KMV model	KMV model outperforms other default prediction models. Captures information from accounting ratios and Moody's credit ratings
Westerlund & Rebeggiani (2012)	Credit rating changes prediction	Merton model	A 12-month lagged DD can predict credit ratings. An asymmetry in predicting down- versus upgrades exist
Afik, Arad & Galil (2015)	Default prediction	Different specifications of the Merton model	Simplified applications of the Merton model outperforms more complex models
Cai, Chen & Dong (2015)	Default prediction	Three different specifications of the Merton model	A naïve DD measure outperforms the other specifications
		Testing Accounting-Based Models	
Pogue & Soldfosky (1968)	Explaining corporate bond ratings	Leverage, Profitability, Size, Earnings stability & EBIT interest coverage	Leverage and profitability have the most significant impact. Overall explanatory power of 56 %
Pinches & Mingo (1973)	Explain industrial bond ratings	Subordination, Years of cons. dividends, Issue size, EBIT interest coverage, Long-term debt ratio & Return on assets	The variables could explain 70 % of the ratings in the original sample, and 60 % in a holdout sample
Gray, Mirkovic & Ragunathan (2006)	Explain corporate credit ratings	Interest coverage, Cash flow, Leverage, Profitability, Systematic risk & Industry concentration	Interest coverage and leverage ratios have the largest impact on credit ratings.
		Horse-Races	
Hillegeist et al. (2004)	Default prediction	Altman's Z-score Ohlson's O-score vs. A variant of the Merton model	The option-based model performs significantly better, even after making adjustments to the Z-score and the O-score
Ghargori, Chan & Faff (2006)	Default prediction	Two option-based models derived from the Merton model vs. An accounting-based model derived from the Z-score and O-score	The option-based models perform similarly and outperform the accounting-based model
Tanthanongsakkun & Treepongkaruna (2008)	Explain corporate credit ratings	Interest coverage & leverage ratios vs. DD, firm size & book-to-market	The option-based model has more explanatory power in explaining credit ratings, 55,25 % vs. 47,47 %

3 Hypotheses

Built upon what was presented in the previous section regarding theories and previous research, this section presents the hypotheses which are to be tested in order to fulfill the purpose of this thesis.

According to Kealhofer and Kurbat (2001), the Merton-like KMV model already captures all default predictive information in Moody's credit ratings. Meanwhile, Bharath and Shumway (2008) show that the Merton model does a good job at capturing the information in the probability estimates from the KMV model. Combining these findings, we find reason to believe that the Merton model will do well in capturing much of the information in the credit ratings from CRAs. Thus we expect, under the assumption that credit ratings are lagged while the DD from the Merton model is instantaneous, the following hypothesis to hold:

H_1: DD from the Merton model has predictive power for future credit rating changes.

Although no previous research has been found testing the predictive power of Altman's Z-score on future corporate credit rating changes, some findings are relevant for our study as they might give some hints as to where our results should point. Previous research has shown that both credit ratings, from the CRAs, and the Altman Z-score have high accuracy ratios in predicting company defaults, and that a firm's history can explain differences in credit ratings. Altman's (1968) suggestions for additional areas of application further indicates that the Z-score could potentially decrease the information asymmetry about credit risk in the market, lowering the cost of monitoring, which is similar to one of the major roles of a credit rating. Given that both are essentially a measure of credit risk and that the Z-scores are up to date, at the moment of the release of inputs needed for calculation, and under the assumption that credit ratings are lagged, we expect to see that the Altman Z-score in our test has predictive power of future corporate credit ratings. Thus, we expect the following hypothesis to hold:

H_2: The Altman Z-score has predictive power for future credit rating changes.

Regarding the relative performance of our models, we expect the Merton model to assess more predictive power and thus stand victorious in the horse-race. Previous research has shown time and again that the option-based models outperform accounting-based models in default prediction. The forward looking aspect of the Merton model theoretically makes the measure better suited for predicting future changes of a lagged variable. We thus expect the DD from the Merton model to have more predictive power than Altman's Z-score.

H_3: DD from the Merton model has more predictive power than Altman's Z-score.

Our final objective is to investigate if there is any asymmetry in the models' abilities in predicting rating up- and downgrades. The majority of previous research, showing stronger market reactions following rating downgrades, indicates that some of the information in rating upgrades is already incorporated in market prices. The findings of Brogaard, Koski and Siegel (2015) on the other hand would lead us to believe that no asymmetry exists. However, following the majority of previous research thus, for the Merton model, lead us to expect that an asymmetry exists and that rating upgrades have more predictive power than downgrades. Surprisingly, Westerlund and Rebeggiani (2012) only found support in the opposite direction. However, due to the lacking sample size of their study, specifically when isolating upgrades from downgrades, we question their result. Regarding Altman's Z-score, there is no clear indication of asymmetry or direction and we do not expect our results to show anything different. We expect the following hypotheses to hold:

H_4: For the Merton model, there is an asymmetry in the predictive power for upgrades and downgrades.

H_5: There is no difference in the predictive power for upgrades and downgrades when using Altman's Z-score.

4 Methodology

This section presents the methodological considerations and approaches of the thesis. This concerns the overall research approach adopted, sample selection, timeframe, data collection and a description of the variables included. Further, an explanation of how regression analyses are to be performed follows before the section is rounded off with a critical review in terms of validity and reliability.

4.1 Research Approach

This thesis aims to test already existing theory, the Merton model and Altman's Z-score, via hypothesis testing on empirical data. Thus, we follow a deductive research approach. This approach means that one, from existing knowledge and theoretical considerations in the area, deduct hypotheses for testing via empirical investigation. Naturally following a deductive research approach is the use of a quantitative research strategy, which enables quantification and analysis of a large amount of data (Bryman & Bell, 2013). From the data obtained, statistical relationships are tested through regression analysis to test the hypotheses.

4.2 Sample and Timeframe

The sample of companies studied in this thesis is all rated non-financial U.S. firms and non-U.S. firms with American Depositary Receipts (ADRs) during the period 2002-2013. In total, the number of companies included in the sample is 1,450, consisting of 1,194 U.S. and 256 non-U.S. firms. In line with previous research; Westerlund and Rebeggiani (2012), Tanthanongsakkun and Treepongkaruna (2008), Bharath and Shumway (2008), financial firms were excluded. The exclusion of financial firms could lead to a sample bias, but was deemed necessary to make because of their unique capital structures and differing accounting standards. Including financial firms would therefore jeopardize the comparability of the firms within the sample.

Including companies with outstanding ADRs implies that some of the companies in the sample are non-U.S. firms, from for example Brazil, Canada, Germany, The Netherlands and the United Kingdom. Including firms from different countries could potentially hamper the comparability of the results. However, when issuing ADRs, firms outside of the U.S. are obliged to follow U.S. accounting standards which indicates that the results should be comparable nevertheless (U.S. Securities and Exchange Commission, 2012).

Furthermore, setting a specific timeframe was necessary to perform this study. Researchers have studied the time-series variation in corporate credit ratings to ascertain shifts in credit rating standards over time. In their article from 1998, Blume, Lim and Mackinlay examined if the corporate credit rating downgrades exceeding rating upgrades of U.S. firms was a result of the CRAs becoming more conservative in assigning ratings. The authors found that credit rating standards had actually become more conservative between 1978 and 1995, concluding that this could partly explain the trend towards lower corporate credit ratings among U.S. firms.

More recent research by Baghai, Servaes and Tamayo (2014) further conclude that credit rating agencies have become more stringent in assigning corporate credit ratings between 1985 and 2009. Alp (2013) found that standards for investment-grade and speculative-grade rated firms tightened and loosened respectively until a breakpoint in 2002, where standards for both categories became more conservative. This followed the corporate failures of for example Enron and WorldCom where credit rating agencies were heavily criticized, as stated earlier, and the introduction of the Sarbanes-Oxley Act, introduced to improve financial disclosure in order to enhance investor protection from accounting frauds.

Therefore, with corporate credit rating standards varying over time, and the major shift in rating standards in 2002, the time period chosen for this study was set to 2002-2013.

4.3 Data

4.3.1 Data Collection

Both credit rating data and the financial data needed for calculations of the Merton model and the Altman Z-score were gathered from Compustat, unless stated otherwise. For comparability reasons, credit ratings and inputs for the Merton model were gathered with yearly frequency, as the Altman Z-score is calculated using accounting data only available on a yearly basis. The credit ratings retrieved from Compustat are the long-term issuer credit ratings from S&P, including the modification of plus and minus signs. To be able to process the rating data and run regressions, letter ratings were translated into numbers, as can be seen in Appendix A, for compatibility with the software program used. As S&P and Moody's are the two largest and most important CRAs (Ghosh, 2013), the choice of which ones' credit ratings to use in our study was based upon availability of credit rating data. To our knowledge, the only possible way to retrieve credit ratings data from Moody's is by manually do it from their homepage. This would be a too tedious process to go through considering the large sample of firms used and the time restrictions of the thesis, thus S&P credit ratings from Compustat were used instead. However, while the CRAs adopt similar rating grades which are treated almost like complete equivalents by market participants, Ghosh (2013) show that Moody's overall assign consistently slightly lower ratings as compared to S&P. The effect differs between industries and is more pronounced regarding the Consumers and Industrials sectors.

4.3.2 Dependent Variables

The dependent variable of interest is the corporate credit rating, more explicitly corporate credit rating *changes*, which can be further separated into rating upgrades and rating downgrades. First, all rating changes in the sample are calculated as the first differences of the translated numerical rating scale. Thereafter, the rating changes are separated into rating upgrades- and downgrades, making up our two dependent variables, and transformed into dummy variables. The rating downgrade dummy variable then takes on the value "1" when a downgrade has occurred and "0" otherwise, while the rating upgrade dummy variable takes on the value "1" when an upgrade has occurred and "0" otherwise.

A drawback of transforming the dependent variables into dummy variables is that credit rating changes of more than one step are disregarded, they are reduced to a one step change. Since 293 rating changes, out of totally 2,853 changes (approximately 10 %), are greater than one step some information is lost that could lessen the potential significance in our regression results. However, as the proportion of rating changes exceeding one step is relatively low we do not think that it will affect our results significantly.

These two dependent variables are further utilized when testing if any asymmetry in the Merton model's and the Altman Z-score's ability to predict rating down- versus upgrades exists.

4.3.3 Independent Variables

The two independent variables examined are the measures from the Merton model and the Altman Z-score model. The variables are firstly constructed as first differences, ΔDD and ΔZ-score, just as for the dependent variables. As we are interested in both models' predictive power of credit rating changes, we use lagged values of both measures. ΔDD and ΔZ-score at an earlier point in time are then paired with later rating changes to ascertain if there is a link between the two measures for a company in earlier time periods and a change in credit rating. For comparability reasons, discussed in section 4.3.1, only a one year lag is included for each independent variable.

The same two independent variables are used when testing for the asymmetry in the models' abilities to predict future rating up- and downgrades.

To account for extreme outliers for the lagged ΔDD and ΔZ-score measures that could skew the results of our regressions a one percent winsorization is performed on the two variables, in a similar manner as done by Bharath and Shumway (2008). When doing this, all observations above the 99^{th} percentile are set to the 99^{th} percentile while the observations below the 1^{st} percentile are set to the 1^{st} percentile. The drawback of doing this procedure is that we do not account for each observation's unique information. However, as these observations do not follow the general pattern and do not seem to fit, the winsorization seems valid to perform in this case. All tables presented in section 5, containing information of the independent variables, are presented after winsorization if not stated otherwise.

4.3.4 Calculations of the Merton Model

The required inputs for calculating DD using the Merton model are; the face value of debt (B), asset value (A_0), asset *growth* volatility (σ_A) and predicted asset growth (μ_A). Because the Merton model applies to market values, of which some are unobservable, contingent claims pricing, through the balance-sheet equivalence between assets and liabilities, have been used to infer the unobserved values from observed values. Preferably, we would follow the complicated iterative procedure introduced by Vassalou and Xing (2004), or alternatively numerically solve the non-linear equation system, as presented in section 2.2, of,

$$E = A_0 N(d_1) - Be^{-rT} N(d_2) \tag{4.1}$$

$$E\sigma_E = A_0 \sigma_A N(d_1) \tag{4.2}$$

as suggested by Gray and Malone (2008) for estimating σ_A and A_0. However, applying any of these methods on a big sample of firms proved to lie outside of our programming capabilities and manually trying to solve for the unobservable σ_A and A_0 using same the methods proved to be far too time-consuming. Instead, a shortcut, as presented by Gray and Malone (2008), was used in order to obtain these values. In this section, a walkthrough of how we obtained the different variables used in the Merton model and how we performed this shortcut is given. The shortcut for estimating A_0 and σ_A is to initially approximate A_0 by E + B and $N(d_1)$ by 1. We then have,

$$A_0 \approx E + B \tag{4.3}$$

which together with equation 4.2 gives us,

$$\sigma_A \approx \frac{E}{E+B} \sigma_E \tag{4.4}$$

Equation 4.3 and 4.4 are used to estimate initial values for A_0 and σ_A. These estimates can be used to calculate initial estimates of current debt (D) through,

$$D = Be^{-rT} - (Be^{-rT} N(-d_2) - A_0 N(-d_1)) \tag{4.5}$$

which is then re-inserted instead of B in equation 4.3 and 4.4 to give closer approximations for A_0 and σ_A. This process can then be repeated and will convert closer and closer to the "true" values of A_0 and σ_A. The above process is performed with two iterations for each firm to obtain our final values. A_0 is then used to calculate the actual μ_A for each year through the return on assets. Lastly, these values are inserted in equation 2.6 to calculate the DD.

The face value of debt was calculated as half of the long-term debt plus short-term debt and current portion of long term debt as suggested Vassalou and Xing (2004), and followed by Bharath och Shumway (2008), and daily market value of equity was gathered from

Datastream in order to calculate σ_E. The 3-month U.S. treasury bill rate of return, and a corresponding measure for the non-U.S. firms, served as a proxy for the risk free interest rate when calculating d_1 and d_2.

4.3.5 Control Variables

To account for other factors being correlated with our dependent variables, whose effect we want to remove, a set of control variables are included in the regressions. This is done in order to clarify the relationship between our dependent and independent variables, without having other variables affecting this relationship. Firstly, to account for the existence of firms from other countries than the U.S. in our sample impacting on the independent variables' prediction abilities of credit rating changes, a country dummy variable is included in the regressions. This dummy variable takes on the value "1" if the firm is from the U.S. and "0" otherwise. The reason for only creating a dummy variable taking into account U.S. versus non-U.S. firms, and not a dummy variable for each country in the sample, is that there are extremely few firms from a single other country than the U.S. as compared to the total number of firms. The country with the second highest number of firms represented in the sample of this thesis is Canada with 51 firms. Creating a dummy variable for Canada alone would lead to a very unbalanced distribution of "1" and "0", and the usefulness of such a variable could therefore be questioned. This problem would be even more pronounced if country dummy variables for other countries, with even less firms represented in the sample, were included in the regressions.

Further, a dummy variable indicating if the firms are investment-graded or speculative-graded is included as a control variable in the regressions. This is done to account for the differences that exist with being investment-grade rated, as compared to being speculative-grade rated, for firms. Being investment-grade rated enable firms access to the whole market, as some investors are prohibited from investing in speculative-grade rated securities. Furthermore, an investment-grade rating enable firms getting both more attractive and flexible covenants and debt issue terms and not to forget a higher status than speculative-grade rated firms (Santos, n.d.).

A variable taking into account the size of the firms is further included in our regressions as a control variable. The rationale underlying the inclusion of firm size is based upon the findings of Vassalou and Xing (2004) who found the variable to be correlated with default risk. The results of their study show the size effect to be most pronounced for firms in the segment with the highest default risk, otherwise the effect were not as significant. Tanthanongsakkun and Treepongkaruna (2008), who include firm size as an explanatory variable based upon Vassalou and Xing's (2004) results, found the variable to be the most significant in explaining credit ratings for Australian firms. The variable is defined in line the abovementioned authors as the market capitalization, and is winsorized using the same procedure as explained in section 4.3.3 to account for extreme outliers.

Lastly, the book-to-market (BM) ratio is further included as a control variable in the regressions, also based on the findings of Vassalou and Xing (2004) and Tanthanongsakkun and Treepongkaruna (2008). They found the variable to be correlated with default risk and able to explain credit ratings respectively. The BM ratio is calculated as the book value of equity divided by the market value of equity, and is also winsorized as explained in section 4.3.3.

4.4 Data Analysis

4.4.1 Regression Analysis

To be able to compare the predictive power of the Merton model and the Altman Z-score, four main regressions are run, one with rating upgrade as dependent variable for each independent variable and one with rating downgrade as dependent variable for each independent variable. The regressions are then analyzed to solve for the ambiguity between the models' relative performance and to find any asymmetry between the predictive power for upgrades and downgrades. These regressions are thereafter rerun, including the control variables mentioned in section 4.3.5, to see if the relationship between the dependent and independent variables is interfered by other variables.

Due to the categorical and non-continuous characteristic of a credit rating, it is deemed a limited dependent variable. As the underlying relationship between a limited dependent variable and the independent variables is non-linear, using a linear regression model is inappropriate (Brooks, 2014). The use of dummy variables as dependent variables enables us using a simplified binary response model. There are basically two different methods that can be used when dealing with binary outcomes, namely the logit and the probit model. As the logit model historically has been favored due to its simplicity (Brooks, 2014) and since most previous researchers have adopted it, the logit model will be used in this thesis and is explained more in detail in the following section.

In general, the choice of which one of the two binary response models, logit or probit, to use is somewhat arbitrary due to the two models' resemblance and both models give virtually indistinguishable results. The only occasion in which the models might give slightly different results is when the distribution of the binary outcomes is very unbalanced. This could be the case for our dependent variables as "1" only occurs around 10 % of the time in our different regressions. To test for the robustness of our choice to use the logit model, we will perform four additional regressions using the probit model as well, in order to ascertain if our results are affected by the functional form of the logit regression model or not. The additional regressions are run with rating downgrades and upgrades as dependent variables for the two independent variables, the one-year lagged ΔDD and ΔZ-score, in total adding up to four regressions.

Eviews is the software tool used when running regressions for this thesis.

4.4.2 The Logit Model

The logit model is based on the cumulative logistic probability distribution function (hence the name of the model) and is defined as follows,

$$F(z_i) = \frac{e^{z_i}}{1+e^{z_i}} = \frac{1}{1+e^{-z_i}} \tag{4.6}$$

where $F(z_i)$ is the cumulative logistic distribution and z_i is the value of the independent variables. The structure of the logit model enables values to lie between 0 and 1 and can therefore be interpreted as probabilities. When estimated, the logit model is as follows,

$$P_i = \frac{1}{1+ e^{-(\alpha+\beta_1 x_2+\cdots+\beta_k x_{ki}+u_i)}} \tag{4.7}$$

where P_i is the probability of y_i=1. When running logit regressions in Eviews the “z-statistic” generated is the equivalent for the “t-statistic” obtained from OLS-regressions. P-values are interpreted as usual for individual coefficient estimates and significance levels and signs can without transformation be compared to OLS-regressions, but the level of the coefficient estimates cannot (Brooks, 2014). The correct way to interpret the coefficient effects on the dependent variable is by looking at their marginal effects (Brooks, 2014), which are calculated using the following function,

$$m_k^{logit} = \beta_k F(z)(1 - F(z)) \tag{4.8}$$

where β_k is the regression coefficient and F(z) is the predicted probability at the means. These marginal effects are not reported automatically by Eviews but have to be calculated manually. When running the logit regressions in Eviews robust standard errors are used (Brooks, 2014).

4.4.3 Regression Interpretation

The intuition is that a rating change can be explained by an increase or decrease in the DD or Z-score. An increase in either the DD or the Z-score indicates a lower probability of default which in turn should result in a better credit rating and vice versa. The marginal effects are interpreted as the increase (decrease) in the probability that the dependent variable is “1” when there is a one unit increase (decrease) in the independent variable.

In our case, the marginal effect shows the probability of an upgrade or downgrade due to a one step increase in either the lagged ΔDD or the lagged ΔZ-score, which in turn means that the DD or Z-score for a firm, one year prior, has increased by one standard deviation or one unit respectively. If the marginal effect for the logit regression is 0.1, with rating upgrade as the dependent variable and lagged ΔDD as the independent variable, the interpretation is that a one standard deviation increase in the DD for a firm increases the likelihood of a rating upgrade, within one year, with 10%.

Regarding the goodness of fit measure, the common ones to examine when dealing with limited dependent variables are the Pseudo-R^2, reported in Eviews as the McFadden R^2, and the percentage correctly predicted. Unlike other R^2 measures, Pseudo-R^2 cannot be interpreted as the percentage variation in the dependent variable explained by the model. However, the value will lie somewhere in between 0 and 1 and the higher the number, the better the fit of the model. Further, Pseudo-R^2's generated when running limited dependent variable regression are low, in general, as compared to other R^2 measures (Brooks, 2014). To be kept in mind is that the Pseudo-R^2 is only comparable for different specifications of the same model (Institute for Digital Research and Education, n.d.) and is therefore only used when examining the predictive power of the models in isolation, not when determining the relative predictive power of the Merton model and Altman's Z-score.

Percentage correctly predicted measures what the name suggests, how many percent of the observations for the independent variable the model correctly predicts, and the higher this number is, the better the fit of the model. Ideally, one would only look at what percentage the model correctly predicts that y=1, but when the sample is unbalanced between "1" and "0", a regression only including an intercept predicting y=1 all the time would outperform any other model. Such a model would, however, not be of any use and instead looking at what percentage the model correctly predicts that both y=1 and y=0 would be more appropriate (Brooks, 2014) which is therefore what we will consider in this thesis. Further, we will also consider the total gain, expressed as percentage points, we get by adding our independent variables to the regressions as compared to what the performance would have been if the regressions would only include an intercept term. When requesting information about percentage correctly predicted in Eviews the proportion of "1" as compared to the total number of observations in the respective regressions will be used as the cutoff point.

The statistical significance, marginal effects and percentage correctly predicted will be examined when determining which of the two models perform better in predicting corporate credit rating changes. The same outputs are looked upon when testing for the potential asymmetry in the models' abilities in predicting downgrades and upgrades. When statistical significance is not achieved for an independent variable, its marginal effect, percentage correctly predicted and Pseudo R^2 will not be further analyzed. Ideally, when making comparisons of the type of regression outputs mentioned above, metrics such as the Bayesian Information Criterion (BIC) or the Accuracy Ratio (AR) could be examined to determine the best fitting model among competing models. Among others, Tanthanongsakkun and Treepongkaruna (2008) and Gharghori, Chan and Faff (2006) utilize these metrics when performing their horse-races. However, due to the complex nature of such metrics, the fact that they are not reported in Eviews and the time-restrictions of the thesis, the comparisons are made by visual examination of the abovementioned regression outputs in accordance with Bongini, Laeven and Majnoni (2002).

4.5 Validity and Reliability

When the quality of a published paper is assessed and evaluated it is often done in terms of validity, both internal and external, and reliability. The internal validity is a measure for the extent to which conclusions regarding causal relationships can be made. Perhaps the most relevant factor concerning this measure is the extent to which our regressions correctly show what we want them to show. Given our somewhat limited knowledge of econometrics and due to not being able to find any previous research to follow, performing exactly or similarly what this thesis is trying to accomplish, there is a possibility that the internal validity is affected by an erroneous methodology. Educated "guesses" made through careful reading of available literature are the foundation for our choice of methodology, but there is still the chance of missing important information that could improve upon the internal validity of this study. Regarding the models used for default prediction however, the accuracy is less questionable. The independent variables were in general defined in accordance with the theories underlying them. The simplifications made concerning the DD measure might slightly deteriorate the usefulness of the variable, but as some needed market values were unobservable, and previous researchers have made similar simplifications, we do not believe this to affect the internal validity to a large extent. As for the dependent variables, we do not believe that disregarding rating changes of more than one step affects the internal validity, as discussed in section 4.3.2. As earlier mentioned, regressions with a set of motivated control variables are run to ensure that the relationship between the dependent and independent variables is not affected by other factors, thus increasing the internal validity. This is further the underlying reason for why robustness checks using the probit model are done. In conclusion, the internal validity is deemed sufficient for a thesis considering the decisions regarding the research approach, variables included and the regressions run.

The external validity, which is a measure of a study's generalizability, in this thesis is deemed high as the sample includes all rated non-financial U.S. firms and firms with ADRs during the period 2002-2013, totaling 1,450 firms. The sample size is, as compared to previous research, quite large and since our sample also includes firms from other countries than the U.S., the external validity should be deemed relatively high. However, one must keep in mind that a clear majority of the firms are from the U.S., just above 80 %, so the generalizability mainly holds within the U.S. The slight differences in the credit ratings assigned by S&P and Moody's respectively, as mentioned in section 4.3.1, indicate that our results might not be entirely generalizable for firms that are by other rating agencies than S&P.

Regarding the reliability, measuring a study's credibility and consistency, we deem it to be somewhat lacking. Although the transparency in the methodology section, regarding sample selection, timeframe and variable- and regression description, enables subsequent research to produce similar results, the conclusions drawn might not entirely the same. Not being able to statistically test the differences when investigating the relative predictive power of the Merton model and Altman Z-score, and the potential asymmetry in the models' abilities in predicting rating down- versus upgrades, is something that could negatively affect the reliability. The conclusions drawn from the analysis of our results are subject to a great deal of subjective judgement.

5 Results and Analysis

The section initially delivers some descriptive statistics concerning the data used in the thesis and is followed by a presentation of the results from our main regressions for rating changes, downgrades and upgrades and the results from robustness checks. Thereafter follows the analysis of our results in connection with the theoretical foundation and previous research.

5.1 Descriptive Statistics

For initial understanding of what types of firms that are included in the thesis, table 5.1 presents the distribution of different rating categories among the 1,450 firms in the sample. The rating category with the highest frequency is BBB and the distribution of investment-grade and speculative-grade rating categories is even, with a slight majority of speculative-grade ratings, 51.84 % versus 48.16 %.

Table 5.1 – Distribution of rating categories.

Rating	# of observations	Rating	# of observations	Rating	# of observations
AAA	109	BBB+	1,495	B	888
AA+	36	BBB	1,953	B-	370
AA	158	BBB-	1,315	CCC+	118
AA-	248	BB+	786	CCC	59
A+	486	BB	1,075	CCC-	11
A	954	BB-	1,261	CC	18
A-	1,060	B+	1,096	C	0
Total	*13,496*				

The sample of 1,450 firms includes a total number of 2,853 observed rating changes out of 13,496 observations[1], evenly distributed between downgrades and upgrades, 1,382 and 1,471 respectively. Due to the nature of first differences and in order to calculate one-year lagged independent variables, all observations for the first two years (2002-2003) are lost when running regressions. Furthermore, data needed for calculating the independent variables are not available for all observations and when running regressions, Eviews only takes into

[1] A rating change is considered observed if the company has a different rating the year prior to the observed year. Out of the 13,496 total rating observations, rating changes cannot be observed for 2002 or for years where the previous year have no available rating data.

account observations that are evenly matched between the dependent and independent variables. Thus, the total number of observations is further reduced and differs slightly between regressions. Due to running several regressions, the exact amount of observations for each regression is not presented here but rather in the respective outputs for convenience.

Calculated DDs and Z-scores, with a one-year lag, for each rating class are shown in table 5.2 to get an indication of the comparability of the different measures. The mean of DD decreases for each drop in rating class and has the highest mean of 20.69 for AAA-rated firms and the lowest mean of 0.875 for CC-rated firms. The mean of Z-score is more volatile and does not drop for each rating class, with the highest mean of 6.41 for AAA-rated firms and the lowest mean of 0.760 for CCC-rated firms.

Table 5.2 – Lagged DD- and Z-score values for different rating categories.

Rating	Lagged DD	Lagged Z-score	Rating	Lagged DD	Lagged Z-score	Rating	Lagged DD	Lagged Z-score
AAA	20.688	6.412	BBB+	8.881	2.089	B	2.924	1.751
AA+	15.703	1.900	BBB	7.917	2.296	B-	1.926	1.535
AA	14.991	2.664	BBB-	6.911	2.537	CCC+	1.801	0.996
AA-	13.217	2.935	BB+	6.053	2.475	CCC	1.438	0.760
A+	12.533	2.937	BB	5.823	2.128	CCC-	0.894	0.941
A	10.506	2.443	BB-	4.467	2.099	CC	0.875	1.684
A-	9.341	2.454	B+	3.596	1.999	C	-	-

The independent variables tested, the one-year lagged ΔDD and ΔZ-score, had, before winsorization, a few extreme max- and minimum values leading to very high standard errors. In table 5.3 some descriptive statistics for the independent variables are presented both before and after winsorization for comparison. The average change in DD and Z-Score, over a year, is 0.382 standard deviations and -0.029 units respectively.

Table 5.3 – Descriptive statistics for the independent variables

	Lagged ΔDD		**Lagged ΔZ-Score**	
	Before winsorization	**After winsorization**	**Before winsorization**	**After winsorization**
Mean	1.283	0.382	0.293	-0.029
Median	0.658	0.658	0.022	0.022
Max	8,942.772	10.050	4,186.517	4.430
Min	-4,873.272	-10.967	-529.837	-5.689
Max- Min	13,816.04	21.017	4,716.354	18.183
Std.dev	113.295	3.536	39.151	1.432

5.2 Regression results

5.2.1 Predicting Future Rating Downgrades

For downgrades, the change in DD shows statistical significance at the 1 % level. The marginal effect is negative, which is to be expected, and a one standard deviation increase (decrease) in the DD decreases (increases) the probability of a rating downgrade within one year by 1 %. The regression correctly predicts if there is a rating downgrade 62.93 % of the time which is a gain of 52.20 percentage points. Introducing the control variables does not change the significance and the coefficient only changes marginally. The regression outputs both with and without control variables and the outputs regarding percentage correctly predicted are presented in Appendix B.

Table 5.4 - Regression for rating downgrade using DD.

Variable	Coefficient	Marginal effect	Std. error	Probability
Intercept	-2.137634	-	0.035040	0.0000
1-year lagged ΔDD	-0.110381	-0,01007	0.008735	0.0000
Dependent variable	*Rating downgrade*			
McFadden R^2	0.0225		**Total obs**	8869
Total percentage correctly predicted	62.93 %		**Obs with dep=1**	951
Total gain	52.20 %		**Obs with dep=0**	7918
Log likelihood	-2953.419			
Avg. log likelihood	-0.333			

For Z-score, it shows statistical significance at the 1 % level when it comes to rating downgrades. Again, the marginal effect is, as expected, negative and a one unit increase (decrease) in the Z-score decreases (increases) the probability of a rating change within one year by 1.46 %. The regression correctly predicts if there is a rating downgrade 65.13 % of the time, a gain of 54.78 percentage points. Adding control variables does not change the significance and the coefficient only changes marginally for Z-score as well. Regression outputs both with and without control variables and the outputs regarding percentage correctly predicted are also presented in Appendix B for the Z-score.

Table 5.5 - Regression for rating downgrade using Z-score.

Variable	Coefficient	Marginal effect	Std. error	Probability
Intercept	-2.175884	-	0.033419	0.0000
1-year lagged ΔZ-score	-0.158582	-0.01457	0.026399	0.0000
Dependent variable	*Rating downgrade*			
McFadden R^2	0.0043		**Total obs**	9838
Total percentage correctly predicted	65.13 %		**Obs with dep=1**	1018
Total gain	54.78 %		**Obs with dep=0**	8820
Log likelihood	-3258.515			
Avg. log likelihood	-0.331			

5.2.2 Predicting Future Rating Upgrades

Turning to the predictive power of changes in the DD on future rating upgrades, the regression shows a statistical significance at the 1 % level. Now, the marginal effect is positive meaning that an increase in the DD, and thus a decrease in the probability of default, results in a higher likelihood of a rating upgrade. A one standard deviation increase in the DD increases the probability of a rating change within one year by 0.08 %. The regression correctly predicts if there is a rating downgrade with 52.87 % which is a gain of 40.75 percentage points. When adding the control variables, the significance remains unchanged and the coefficient only changes marginally. The regression outputs both with and without control variables and the outputs regarding percentage correctly predicted are presented in Appendix C.

Table 5.6 - Regression for rating upgrade using DD.

Variable	Coefficient	Marginal effect	Std. error	Probability
Intercept	-2.035198	-	0.033735	0.0000
1-year lagged ΔDD	0.075874	0.00793	0.008660	0.0000
Dependent variable	*Rating upgrade*			
McFadden R^2	0.0095		**Total obs**	8869
Total percentage correctly predicted	52.87 %		**Obs with dep=1**	1075
Total gain	40.75 %		**Obs with dep=0**	7794
Log likelihood	-3244.282			
Avg. log likelihood	-0.366			

For Z-score, the significance is also at the 1 % level and the marginal effect is positive. A one unit increase (decrease) in the Z-score increases (decreases) the probability of a rating upgrade within one year by 2.22 %. The regression correctly predicts if there is a rating downgrade 55.10 %, a gain of 43.07 percentage points. Checking for the changed effect when including size, investment or speculative grade and country, has no impact on the significance and the coefficient only changes marginally. Regression outputs both with and without control variables and the outputs regarding percentage correctly predicted are also presented in Appendix C for the Z-score.

Table 5.7 - Regression for rating upgrade using Z-score.

Variable	Coefficient	Marginal effect	Std. error	Probability
Intercept	-1.998877	-	0.031331	0.0000
1-year lagged ΔZ-score	-0.209049	0.02207	0.033927	0.0000
Dependent variable	*Rating upgrade*			
McFadden R^2	0.0062		**Total obs**	9838
Total percentage correctly predicted	55.10 %		**Obs with dep=1**	1184
Total gain %	43.07 %		**Obs with dep=0**	8654
Log likelihood	-3594.180			
Avg. log likelihood	-0.365			

5.2.3 Robustness Checks with the Probit Model

In line with what was said in section 4.4.1, regarding which of the two binary response models, logit or probit, to use when dealing with an unbalanced dummy variable, four regressions were run in order to ascertain if the functional form of the logit model affects our results. The results of these regressions are presented in Appendix D. In general, the outputs of the regression models with the same specifications show very small differences in statistical significance, coefficient values and standard errors. There is no change in the statistical significance for the independent variable for any of the six regressions compared. A trend of slightly larger absolute values, although small differences, for both coefficients and standard errors can be seen when running logit regressions. This is probably due to the logit and probit model using different distributions, the cumulative logistic probability- versus the cumulative normal distribution. However, as the differences when comparing the McFadden

R^2 and percentage correctly predicted outputs, presented for the probit regressions in Appendix D, are almost indistinguishable we conclude that the functional form of the logit model is not affecting the results of our regressions.

5.3 Result Analysis

The empirical results are analyzed in connection to the theoretical foundation and previous research. For convenience, a result summary is presented in table 5.8 covering the most considered outputs for the analysis.

Table 5.8 – Summary of regression outputs for analysis.

Regression	Variable	Probability	Marginal effect	Total percentage correctly predicted	McFadden R^2
Rating downgrade	1-year lagged ΔDD	0.0000	-0,01007	62.93 %	0.0225
	1-year lagged ΔZ-score	0.0000	-0.01457	65.13 %	0.0043
Rating upgrade	1-year lagged ΔDD	0.0000	0.00793	52.87 %	0.0095
	1-year lagged ΔZ-score	0.0000	0.02207	55.10 %	0.0062

The results show that our independent variables are statistically significant, at the 1 % level, for both upgrades and downgrades and thus that the measures do have predictive power for overall rating changes one year ahead, supporting both H_1 and H_2. Further, the signs of the marginal effects are also in line with previous research and the theoretical view that an increase in either DD or Z-score indicates a lower probability of default which in turn should result in a better credit rating and vice versa. Due to the use of lagged variables, both measures now prove to already contain some of the information presented by a credit rating one year later, supporting the view that credit ratings are lagged.

"Some of the information", however, is quite vague and the Merton model can *only* generate correct predictions, i.e. predict if a firm will or will not receive a rating upgrade or downgrade, about 53 % and 63 % of the time respectively and the marginal effects and

Pseudo-R^2s are extremely low for all regressions[2]. Thus, although statistically significant, changes in DD do not seem to provide sufficient information for predicting future rating changes and other variables must be included for more accuracy. The result is similar with previous research from Bharath and Shumway (2008), although applied on lagged rating changes rather than company defaults, in that the DD calculated by the Merton model provides useful guidance for building prediction models. Thus, the result does not support that the Merton model does a good job in capturing much of the information in the credit ratings from the CRAs, as reasoned by combining results from Kealhofer and Kurbat (2001) and Bharath and Shumway (2008) in section 3. Notable is that Kealhofer and Kurbat (2001) used ratings from Moody's while S&P ratings are used in this study, which might not be entirely interchangeable, as discussed in section 4.3.1. For the Z-score, a similar line of reasoning can be had. Similar percentage correctly predicted outcomes imply that much of the variation of the independent variables remains unexplained and that a changed Z-score is only a piece of the puzzle.

Comparing the results of our models' relative performances, the statistical significances are indistinguishable but Z-score seems to produce slightly higher percentage correctly predicted outcomes and stronger, although still very weak, marginal effects, and the result does thus not support H_3. Interestingly, the Pseudo-R^2s show the opposite relationship but will not be further analyzed in the comparison. However, the differences between the models are not big enough to proclaim any clear winner[3]. The relative performances, applied on predicting future rating changes, are different from previous research in the area of default prediction and credit rating prediction presented in section 2.4 and is not intuitive given the forward-looking

[2] We say *only* even though the percentages correctly predicted outcomes do not deviate much, in terms of interpretation, from previous research for explaining corporate credit ratings. For comparison, although the measures are far from entirely interchangeable, the explanatory power using the Merton model in combination with firm size and book-to-market was found to be around 55% by Tanthanongsakkun and Treepongkaruna (2008). Further, for default prediction, Du and Suo (2007) showed correctly predicted outcomes of around 50 % and 40 % using a reduced form of the Merton model and the original respectively. Either way, previous research is conclusive in that there remains a lot to be explained which is similar to our result.

[3] The deciding rules, for if there is enough difference in the models' relative performances, to proclaim one model to be better than the other are subjective and somewhat vague as discussed in section 4.5. It is our interpretation that the measures are close enough to not being able to draw any conclusions in this aspect.

aspects of DD as compared to the backwards-looking aspect of the Z-score[4]. One factor contributing to this result could be that the DD-measure is forced out of its element by the necessity of only looking at one year lags, due to comparability reasons, and that its potential for investigating several shorter lag periods is overlooked. It is possible that further research would show stronger predictive power using different time lags for DD while Z-score would be hard to improve upon. Another possible reason for the result could be due to the simplifications made in order to calculate the DD, and that following Vassalou and Xing's (2004) iterative procedure would improve on the measure, however unlikely due to the bulk of previous research supporting simpler versions of the model. For now, we can only conclude that no model clearly outperforms the other when using a one year lag.

Showing predictive power for both upgrades and downgrades using DD differentiates our results to the findings of Westerlund and Rebeggiani (2012) and could partly be due to a larger sample size. The conclusion, however, is the same and the results support H_4. An asymmetry exists with downgrades being more accurately predicted than upgrades. Further, the marginal effect for downgrades is also stronger, although slightly so, and the Pseudo-R^2 is stronger but hard to interpret. The result is surprising and does not support our view that upgrades should be more easily predicted as reasoned in section 3 due to previous research showing weaker market reactions to rating upgrades. The result could instead be explained by arguing that the interdependency between rating agencies and issuers incites the rating agencies to think twice before downgrading a company as reasoned in section 1.3.1. Rating agencies have been criticized for untimely rating changes and the main focus has not been untimely upgrades. Perhaps downgrades are more lagged in general which could be a possible explanation for our results. For Z-score, the percentage correctly predicted outcomes are also higher for downgrades but the marginal effect is weaker, and so is the Pseudo-R^2, making the interpretation less clear. With the measures pointing in different directions we cannot really say that either upgrades or downgrades are more easily predicted which was expected

---

[4] Notable here is that while one could argue that the result to some extent supports the findings of Pogue and Soldofsky (1968) in that a firms history better explain differences in corporate bond ratings rather than future prospects, however it is important to keep in mind that the article was published before the Merton model was created and the comparison is thus not really relevant.

(although the sprawling outputs were not) and supports H_5 due to no clear indication of asymmetry regarding the Z-score.

Overall, the discrepancy between if a rating change occur and if it was predicted by either DD or Z-Score, combined with the low marginal effects and Pseudo-R^2s, can have many possible explanations. An important factor could be the chosen time lag, which was chosen for comparability reasons between our models. Our findings, that the predictive power is relatively weak for both measures, only concludes that it is weak when using a one year lag which is a rather long period. Also, as discussed in section 2.1, the definition of credit risk by S&P is not only a firm's ability but also its *willingness* to fulfill its financial obligations and the credit ratings further differ from our independent variables in that they are supposedly long-term looking and based on qualitative aspects. Furthermore, there is also the possibility that changes in DD and Z-score in fact should have led to a rating change for a specific observation but does not due to credit ratings not being a flawless measure of credit risk, potentially causing weaker results than if testing our models on a hypothetically true measure of credit risk.

6 Conclusions and Future Research

In this final section, conclusions are made and hypotheses are answered, based on the analysis, in connection to the purpose of this thesis. Lastly, suggestions for previous research are presented.

This study is the first to investigate the relative performance of the Merton model and Altman's Z-score in predicting future corporate credit rating changes, by conducting a horse-race between the two models, and to find if any asymmetry between upgrades or downgrades exists. Using logit regressions, on a sample of 1,450 firms, both models prove to have some predictive power for rating changes one year ahead but the goodness of fit is mediocre and the marginal effects are low. Although the Z-score show slightly better results, in terms of percentage correctly predicted outcomes, it is concluded that no clear difference in the relative performance can be found. Meanwhile, it is concluded that the Merton model has more predictive power for downgrades than for upgrades while no such asymmetry can be found for the Z-score.

Finding existing, although weak, predictive power for corporate credit rating changes within one year, using basic models such as the Merton model and Altman's Z-score, has several implications.

Firstly, since quantitative models of credit risk can predict a changed credit rating to some extent it extends our knowledge, and supports the criticism, that rating changes from the CRAs are untimely in reflecting a subject's changed credit risk. Lagged credit ratings could lead to incorrectly rated firms suffering, or gaining, from financing costs not appropriately reflecting the firms risk level. Further, among other things, CFOs could make faulty decisions concerning debt-levels and investors could make bad investment decisions based on faulty information.

Secondly, while this study offers limited contribution in showing how great this lag is due to limitations in our data, for comparability reasons, the result still underlines the importance of continued curiosity in the area and could serve as a benchmark for subsequent research. Although both models show limited power in accurately predicting rating changes, Altman's Z-score and other accounting-based models lack the ability of incorporating instantaneously updated information and are thus hard to improve upon. The Merton model and other option-based models, although more complex to use, are not hindered by such restrictions. We suggest that dropping the accounting-based models and instead focusing on investigating the Merton model and a plentitude of alternative option-based models, while incorporating more timely precise data for rating changes and including several different lag-periods, is appropriate for building forecasting models of better accuracy and precision.

Lastly, until more accurate prediction models are conducted, or rendered inadequate, through future research in the area, corporate credit ratings still seemingly carry unique information that is not totally embedded in market prices, or concealed within the books, and thus still have an important role to play in decreasing the asymmetry between issuers and creditors.

Further, that downgrades are more easily predicted, using a model which incorporates market prices for its calculations, is not entirely intuitive and is not in line with previous research but could support that the interdependency between CRAs and issuers affect the ratings. It could imply that downgrades are, to a greater extent, affected by the untimeliness of rating changes and that the CRAs more rapidly adjust the ratings of firms that decrease their credit risk while actual increases in credit risk are adjusted for more slowly. Thus, rating changes finally revealing information regarding the worsened conditions might not be handed out until things have gotten bad enough, and thus possibly carry information of greater vicissitude, which could explain stronger market reactions following downgrades. However, this remains speculative until further research can find evidence for different quantity of information between upgrades and downgrades, or show that rating changes in the form of downgrades more often occur in greater steps than for upgrades.

References

Association for Financial Professionals. (2002). Rating Agencies Survey: Accuracy, Timeliness, and Regulation. Available online: http://www.worldbusinesslaw.net/Gnosys/ratings_survey.pdf [Accessed 15 May 2016]

Afik, Z., Arad, O. & Galil, K. (2015). Using Merton model for default prediction: An empirical assessment of selected alternatives, *Journal of Empirical Finance*, vol.35, p. 43-67. Available online: LUBsearch [Accessed 6 April 2016]

Alp, A. (2013). Structural Shifts in Credit Rating Standards, *Journal of Finance*, vol. 68, no. 6, pp. 2435-2470. Available online: LUBsearch [Accessed 8 April 2016]

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, vol. 23, no. 4, pp. 589-609. Available online: LUBsearch [Accessed 14 April 2016]

Amato, J. D. & Furfine, C. H. (2004). Are Credit Ratings Procyclical? *Journal of Banking & Finance*, vol. 28, no. 11, pp. 2641-2677. Available online: LUBsearch [Accessed 12 May 2016]

Baghai, R. P., Servaes, H. & Tamayo, A. (2014). Have Rating Agencies Become More Conservative? Implications for Capital Structure and Debt Pricing, *Journal of Finance*, vol. 69, no. 5, pp. 1961-2005. Available online: LUBsearch [Accessed 8 April 2016]

Beaver, W. H. (1996). Financial Ratios as Predictors of Failure, *Journal of Accounting Research*, vol. 4, pp. 71-111. Available online: LUBsearch [Accessed 14 April 2016]

Bharath, S. T. & Shumway, T. (2008). Forecasting Default with the Merton Distance to Default, *The Review of Financial Studies*, vol. 21, no. 3, pp. 1339-1369. Available online: LUBsearch [Accessed 8 April 2016]

Blume, M. E., Lim, F. & Mackinlay, A. C. (1998). The Declining Credit Quality of U.S. Corporate Debt: Myth or Reality?, *Journal of Finance*, vol. 53, no. 4, pp. 1389-1413. Available online: LUBsearch [Accessed 8 April 2016]

Bongini, P., Laeven, L. & Majnoni, G. (2002). How Good is the Market at Assessing Bank Fragility? A Horse Race Between Different Indicators, *Journal of Banking and Finance*, vol. 26, pp. 1011-1028. Available online: LUBsearch [Accessed 20 May 2016]

Brogaard, J., Koski, J. L. & Siegel, A. F. (2015). The Information Content of Credit Rating Changes: Evidence from Trading Volume. Available online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2619522 [Accessed 7 April 2016]

Brooks, C. (2014). Introductory Econometrics for Finance, 3rd edn, Cambridge: Cambridge University Press.

Bryman, A. & Bell, E. (2013). Företagsekonomiska forskningsmetoder, Stockholm: Liber AB.

Cai, CH. W., Chen, Y. L. & Dong, X. J. (2015). Forecasting ability of alternative distance-to-default based on the Merton model, *Management, Information and Educational Engineering*, pp. 85-92. Available online: http://www.crcnetbase.com/doi/pdf/10.1201/b18558-19 [Accessed 8 April 2016]

Charitou, D., Dionysiou, D., Lambertides, N. & Trigeorgis, L. (2013). Alternative Bankruptcy Prediction Models Using Option-Pricing Theory, *Journal of Banking and Finance*, vol. 37, pp. 2329-2341. Available online: LUBsearch [Accessed 6 April 2016]

Dichev, I. D. & Piotroski, J. D. (2001). The Long-Run Stock Returns Following Bond Ratings Changes, *Journal of Finance*, vol. 56, no. 1, pp. 173-203. Available online: LUBsearch [Accessed 7 April 2016]

Du, Y. & Suo, W. (2007). Assessing Credit Quality from the Equity Market: Can a Structural Approach Forecast Credit Ratings?, *Canadian Journal of Administrative Sciences*, vol. 24, no. 3, pp. 212-228. Available online: LUBsearch [Accessed 8 April 2016]

Feinberg, M., Shelor, R. & Jiang, J. (2004). The Effect of Solicitation and Independence on Corporate Bond Ratings, *Journal of Business Finance & Accounting*, vol. 31, no. 9, pp. 1327-1353. Available online: LUBsearch [Accessed 12 May 2016]

Freitas, A. & Minardi, A. M. (2013). The Impact of Credit Rating Changes in Latin American Stock Markets, *Brazilian Administration Review*, vol. 10, no. 4, pp. 439-461. Available online: LUBsearch [Accessed 7 April 2016]

Ganguin, B. & Bilardello, J. (2005). Fundamentals of Corporate Credit Analysis. New York: McGraw-Hill.

Gharghori, P., Chan, H. & Faff, R. (2006). Investigating the Performance of Alternative Default-Risk Models: Option-Based Versus Accounting-Based Approaches, *Australian Journal of Management*, vol. 31, no. 2, pp. 207-234. Available online: LUBsearch [Accessed 29 March 2016]

Ghosh, S. (2013). A Study of Differences in Standard and Poor's and Moody's Corporate Credit Ratings, working paper, The Leonard N. Stern School of Business Glucksman Institute for Research in Securities Markets. Available online: http://www.stern.nyu.edu/sites/default/files/assets/documents/con_041250.pdf [Accessed 7 April 2016]

Graham, J. R. & Harvey, C. R. (2001). The Theory and Practice of Corporate Finance: Evidence From the Field, *Journal of Financial Economics*, vol. 60, no. 2, pp. 187-243. Available online: LUBsearch [Accessed 13 April 2016]

Gray, S., Mirkovic, A. & Ragunathan, V. (2006). The Determinants of Credit Ratings: Australian Evidence, *Australian Journal of Management*, vol. 31, no. 2, pp. 333-354. Available online: LUBsearch [Accessed 13 April 2016]

Gray, D. & Malone, S. W. (2008). Macrofinancial Risk Analysis. Chichester: Wiley.

Hand, J. R. M., Holthausen, R. W. & Leftwich, R. W. (1992). The Effect of Bond Rating Agency Announcements on Bond and Stock Prices, *Journal of Finance*, vol. 47, no. 2, pp. 733-752. Available online: LUBsearch [Accessed 7 April 2016]

Hillegeist, S.A., Keating, E. K., Cram, D. P. & Lundstedt, K. G. (2004). Assessing the Probability of Bankruptcy, *Review of Accounting Studies*, vol. 9, no. 1, pp. 5-34. Available online: LUBsearch [Accessed 14 April 2016]

Institute for Digital Research and Education. (n.d.). FAQ: What are pseudo R-squareds? Available online: http://www.ats.ucla.edu/stat/mult_pkg/faq/general/Psuedo_RSquareds.htm [Accessed 25 May 2016]

Kealhofer, S. & Kurbat, M. (2001). The Default Prediction Power of the Merton Approach, Relative to Debt Ratings and Accounting Variables, KMV LLC. Available online: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=6936BCA6636D5133B152D3A5E1BB14CC?doi=10.1.1.101.7279&rep=rep1&type=pdf [Accessed 8 April 2016]

Langohr, H. & Langohr, P. (2009). The Rating Agencies and Their Credit Ratings: What They Are, How They Work, and Why They are Relevant, Chichester: John Wiley & Sons.

Merton, R. E. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, vol. 29, no. 2, pp. 449-470. Available online: LUBsearch [Accessed 14 April 2016]

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, vol. 1, no. 1, pp. 109-131. Available online: LUBsearch [Accessed 14 April 2016]

Orth, W. (2012). The Predictive Accuracy of Credit Ratings: Measurement and Statistical Inference, *International Journal of Forecasting,* vol. 28, no. 1, pp. 288-296. Available online: LUBsearch [Accessed 12 May 2016]

Pinches, G. E. & Mingo, K. A. (1973). A Multivariate Analysis of Industrial Bond Rating, *Journal of Finance*, vol. 28, no. 1, pp. 1-18. Available online: LUBsearch [Accessed 19 April 2016]

Pogue, T. F. & Soldofsky, R. M. (1969). What's in a Bond Rating, *The Journal of Financial and Quantitative Analysis*, vol. 4, no. 2, pp. 201-228. Available online: LUBsearch [Accessed 19 April 2016]

Santos, K. (n.d.). Corporate Credit Ratings: A Quick Guide. Available online: https://www.treasurers.org/ACTmedia/ITCCMFcorpcreditguide.pdf [Accessed 19 May 2016]

Standard and Poor's (2014a). Guide to Credit Rating Essentials. Available online: http://www.spratings.com/documents/20184/760102/SPRS_Understanding-Ratings_GRE.pdf/298e606f-ce5b-4ece-9076-66810cd9b6aa [Accessed 25 April 2016]

Standard and Poor's (2014b). Corporate Ratings Methodology. Available online: http://www.spratings.com/documents/20184/774196/Corporate+Ratings+Methodology.pdf/8fd4392a-4aae-4669-bd74-a9b86e18d781 [Accessed 25 April 2016]

Standard and Poor's (2016). Understanding Ratings. The rating process. Available online: http://www.spratings.com/en_US/understanding-ratings#thirdPage [Accessed 25 April 2016]

Tanthanongsakkun, S. & Treepongkaruna, S. (2008). Explaining Credit Ratings of Australian Companies - An Application of the Merton Model, *Australian Journal of Management*, vol. 33, no. 2, pp. 261-276. Available online: LUBsearch [Accessed 28 March 2016]

U.S. Securities Exchange Commision. (2012). Investor Bulletin: American Depository Receipts. Available online: https://www.sec.gov/investor/alerts/adr-bulletin.pdf [Accessed 10 May 2016]

Vassalou, M. & Xing, Y. (2004). Default Risk in Equity Returns, *Journal of Finance*, vol. 59, no. 2, pp. 831-868. Available online: LUBsearch [Accessed 10 May 2016]

Westerlund, M. & Rebeggiani, S. (2012). Rating Changes - Can They Be Predicted with the Merton Model? Master Thesis, Department of Business Administration, Lund University. Available online: https://lup.lub.lu.se/search/publication/2540753 [Accessed 18 April 2016]

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models, *Journal of Accounting Research,* vol. 22, pp. 59-82. Available online: LUBsearch [Accessed 14 April 2016]

Appendix A

Presented below is a translation table of letter credit rating into a numerical rating scale.

Appendix A.1 – Translation of letter ratings into a numerical rating scale.

S&P rating	Ordinal scale	S&P rating	Ordinal scale	S&P rating	Ordinal scale
AAA	1	BBB+	8	B	15
AA+	2	BBB	9	B-	16
AA	3	BBB-	10	CCC+	17
AA-	4	BB+	11	CCC	18
A+	5	BB	12	CCC-	19
A	6	BB-	13	CC	20
A-	7	B+	14	C	21

Appendix B

Presented below are the outputs from the regressions made with rating downgrades as dependent variable. First, the regression output with only the independent variable of interest is presented followed by the regression output including the control variables. Lastly, the output regarding the percentage correctly predicted is presented. This is done firstly for the measure from the Merton model and secondly for the Altman Z-score.

Appendix B.1 - Regression for rating downgrade using DD.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/20/16 Time: 11:18
Sample (adjusted): 3 17400
Included observations: 8869 after adjustments
Convergence achieved after 5 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.137634	0.035040	-61.00581	0.0000
LAG1DD_WINSORIZE...	-0.110381	0.008735	-12.63647	0.0000

McFadden R-squared	0.022526	Mean dependent var	0.107227
S.D. dependent var	0.309420	S.E. of regression	0.307301
Akaike info criterion	0.666461	Sum squared resid	837.3463
Schwarz criterion	0.668059	Log likelihood	-2953.419
Hannan-Quinn criter.	0.667005	Deviance	5906.839
Restr. deviance	6042.965	Restr. log likelihood	-3021.482
LR statistic	136.1259	Avg. log likelihood	-0.333005
Prob(LR statistic)	0.000000		
Obs with Dep=0	7918	Total obs	8869
Obs with Dep=1	951		

Appendix B.2 - Regression for rating downgrade using DD and control variables.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/23/16 Time: 11:43
Sample (adjusted): 3 17399
Included observations: 8006 after adjustments
Convergence achieved after 6 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.856961	0.114025	-16.28555	0.0000
LAG1DD_WINSORIZED	-0.109809	0.010032	-10.94578	0.0000
BM_WINSORIZED	0.434090	0.077379	5.609909	0.0000
COUNTRY	-0.251944	0.090306	-2.789873	0.0053
INV_GRADE_VS_SPEC_GRAD...	-0.500021	0.080575	-6.205679	0.0000
SIZE_WINSORIZED	-3.61E-06	1.72E-06	-2.102638	0.0355

McFadden R-squared	0.042827	Mean dependent var	0.110917
S.D. dependent var	0.314049	S.E. of regression	0.308355
Akaike info criterion	0.668510	Sum squared resid	760.6617
Schwarz criterion	0.673747	Log likelihood	-2670.047
Hannan-Quinn criter.	0.670303	Deviance	5340.094
Restr. deviance	5579.027	Restr. log likelihood	-2789.514
LR statistic	238.9330	Avg. log likelihood	-0.333506
Prob(LR statistic)	0.000000		
Obs with Dep=0	7118	Total obs	8006
Obs with Dep=1	888		

Appendix B.3 – Percentage correctly predicted rating downgrades using only DD.

Expectation-Prediction Evaluation for Binary Specification
Equation: DOWNGRADE_DD
Date: 05/20/16 Time: 13:28
Success cutoff: C = 0.1072

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	5057	427	5484	0	0	0
P(Dep=1)>C	2861	524	3385	7918	951	8869
Total	7918	951	8869	7918	951	8869
Correct	5057	524	5581	0	951	951
% Correct	63.87	55.10	62.93	0.00	100.00	10.72
% Incorrect	36.13	44.90	37.07	100.00	0.00	89.28
Total Gain*	63.87	-44.90	52.20			
Percent Gain...	63.87	NA	58.47			

Appendix B.4 - Regression for rating downgrade using Z-score.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/18/16 Time: 16:08
Sample (adjusted): 3 17400
Included observations: 9838 after adjustments
Convergence achieved after 5 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.175884	0.033419	-65.10979	0.0000
LAG1ZSCORE_WINSORIZE...	-0.158582	0.026399	-6.007105	0.0000

McFadden R-squared	0.004322	Mean dependent var	0.103476
S.D. dependent var	0.304595	S.E. of regression	0.304258
Akaike info criterion	0.662841	Sum squared resid	910.5484
Schwarz criterion	0.664303	Log likelihood	-3258.515
Hannan-Quinn criter.	0.663336	Deviance	6517.029
Restr. deviance	6545.315	Restr. log likelihood	-3272.658
LR statistic	28.28608	Avg. log likelihood	-0.331217
Prob(LR statistic)	0.000000		

Obs with Dep=0	8820	Total obs	9838
Obs with Dep=1	1018		

Appendix B.5 - Regression for rating downgrade using Z-score and control variables.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/23/16 Time: 11:46
Sample (adjusted): 3 17399
Included observations: 7442 after adjustments
Convergence achieved after 6 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.940920	0.117401	-16.53233	0.0000
LAG1ZSCORE_WINSORIZED	-0.185320	0.028160	-6.581069	0.0000
BM_WINSORIZED	0.483784	0.082580	5.858379	0.0000
COUNTRY	-0.279954	0.091257	-3.067764	0.0022
INV_GRADE_VS_SPEC_GRAD...	-0.492149	0.085081	-5.784514	0.0000
SIZE_WINSORIZED	-2.54E-06	1.76E-06	-1.438492	0.1503

McFadden R-squared	0.029755	Mean dependent var	0.105482
S.D. dependent var	0.307195	S.E. of regression	0.303497
Akaike info criterion	0.655489	Sum squared resid	684.9354
Schwarz criterion	0.661064	Log likelihood	-2433.075
Hannan-Quinn criter.	0.657404	Deviance	4866.149
Restr. deviance	5015.382	Restr. log likelihood	-2507.691
LR statistic	149.2334	Avg. log likelihood	-0.326938
Prob(LR statistic)	0.000000		

Obs with Dep=0	6657	Total obs	7442
Obs with Dep=1	785		

Appendix B.6 – Percentage correctly predicted rating downgrades using only Z-score.

Expectation-Prediction Evaluation for Binary Specification
Equation: DOWNGRADE_ZSCORE
Date: 05/20/16 Time: 13:29
Success cutoff: C = 0.10347

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	5903	514	6417	0	0	0
P(Dep=1)>C	2917	504	3421	8820	1018	9838
Total	8820	1018	9838	8820	1018	9838
Correct	5903	504	6407	0	1018	1018
% Correct	66.93	49.51	65.13	0.00	100.00	10.35
% Incorrect	33.07	50.49	34.87	100.00	0.00	89.65
Total Gain*	66.93	-50.49	54.78			
Percent Gain...	66.93	NA	61.10			

Appendix C

Presented below are the outputs from the regressions made with rating upgrades as dependent variable. First, the regression output with only the independent variable of interest is presented followed by the regression output including the control variables. Lastly, the output regarding the percentage correctly predicted is presented. This is done firstly for the measure from the Merton model and secondly for the Altman Z-score.

Appendix C.1 - Regression for rating upgrade using DD.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/20/16 Time: 11:19
Sample (adjusted): 3 17400
Included observations: 8869 after adjustments
Convergence achieved after 5 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.035198	0.033735	-60.32821	0.0000
LAG1DD_WINSORIZE...	0.075874	0.008660	8.761756	0.0000

McFadden R-squared	0.009548	Mean dependent var	0.121209
S.D. dependent var	0.326388	S.E. of regression	0.325511
Akaike info criterion	0.732051	Sum squared resid	939.5227
Schwarz criterion	0.733650	Log likelihood	-3244.282
Hannan-Quinn criter.	0.732596	Deviance	6488.564
Restr. deviance	6551.111	Restr. log likelihood	-3275.555
LR statistic	62.54733	Avg. log likelihood	-0.365800
Prob(LR statistic)	0.000000		
Obs with Dep=0	7794	Total obs	8869
Obs with Dep=1	1075		

Appendix C.2 - Regression for rating upgrade using DD and control variables.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/23/16 Time: 11:48
Sample (adjusted): 3 17399
Included observations: 8006 after adjustments
Convergence achieved after 6 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.359041	0.103111	-13.18032	0.0000
LAG1DD_WINSORIZED	0.078778	0.010079	7.815766	0.0000
BM_WINSORIZED	-0.373024	0.067408	-5.533791	0.0000
COUNTRY	-0.176103	0.089704	-1.963156	0.0496
INV_GRADE_VS_SPEC_GRAD...	-0.634757	0.078252	-8.111694	0.0000
SIZE_WINSORIZED	-7.62E-07	1.27E-06	-0.599885	0.5486

McFadden R-squared	0.026302	Mean dependent var	0.118286
S.D. dependent var	0.322967	S.E. of regression	0.319949
Akaike info criterion	0.709371	Sum squared resid	818.9410
Schwarz criterion	0.714608	Log likelihood	-2833.612
Hannan-Quinn criter.	0.711163	Deviance	5667.224
Restr. deviance	5820.307	Restr. log likelihood	-2910.154
LR statistic	153.0833	Avg. log likelihood	-0.353936
Prob(LR statistic)	0.000000		
Obs with Dep=0	7059	Total obs	8006
Obs with Dep=1	947		

Appendix C.3 – Percentage correctly predicted rating upgrades using only DD.

Expectation-Prediction Evaluation for Binary Specification
Equation: UPGRADE_DD
Date: 05/20/16 Time: 13:29
Success cutoff: C = 0.1212

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	4062	448	4510	0	0	0
P(Dep=1)>C	3732	627	4359	7794	1075	8869
Total	7794	1075	8869	7794	1075	8869
Correct	4062	627	4689	0	1075	1075
% Correct	52.12	58.33	52.87	0.00	100.00	12.12
% Incorrect	47.88	41.67	47.13	100.00	0.00	87.88
Total Gain*	52.12	-41.67	40.75			
Percent Gain...	52.12	NA	46.37			

Appendix C.4 - Regression for rating upgrade using Z-score.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/18/16 Time: 16:10
Sample (adjusted): 3 17400
Included observations: 9838 after adjustments
Convergence achieved after 5 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.998877	0.031331	-63.79780	0.0000
LAG1ZSCORE_WINSORIZE...	0.209049	0.033927	6.161627	0.0000

McFadden R-squared	0.006215	Mean dependent var	0.120350
S.D. dependent var	0.325387	S.E. of regression	0.324611
Akaike info criterion	0.731079	Sum squared resid	1036.442
Schwarz criterion	0.732542	Log likelihood	-3594.180
Hannan-Quinn criter.	0.731575	Deviance	7188.359
Restr. deviance	7233.313	Restr. log likelihood	-3616.656
LR statistic	44.95361	Avg. log likelihood	-0.365336
Prob(LR statistic)	0.000000		

Obs with Dep=0	8654	Total obs	9838
Obs with Dep=1	1184		

Appendix C.5 - Regression for rating upgrade using Z-score and control variables.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 05/23/16 Time: 11:52
Sample (adjusted): 3 17399
Included observations: 7442 after adjustments
Convergence achieved after 6 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.416627	0.104412	-13.56762	0.0000
LAG1ZSCORE_WINSORIZED	0.223369	0.036024	6.200548	0.0000
BM_WINSORIZED	-0.355482	0.066132	-5.375350	0.0000
COUNTRY	-0.098240	0.091210	-1.077078	0.2814
INV_GRADE_VS_SPEC_GRAD...	-0.587415	0.079834	-7.357923	0.0000
SIZE_WINSORIZED	4.52E-07	1.26E-06	0.359213	0.7194

McFadden R-squared	0.023001	Mean dependent var	0.122682
S.D. dependent var	0.328094	S.E. of regression	0.325211
Akaike info criterion	0.728959	Sum squared resid	786.4473
Schwarz criterion	0.734534	Log likelihood	-2706.456
Hannan-Quinn criter.	0.730874	Deviance	5412.911
Restr. deviance	5540.346	Restr. log likelihood	-2770.173
LR statistic	127.4342	Avg. log likelihood	-0.363673
Prob(LR statistic)	0.000000		

Obs with Dep=0	6529	Total obs	7442
Obs with Dep=1	913		

Appendix C.6 – Percentage correctly predicted rating upgrades using only Z-score.

Expectation-Prediction Evaluation for Binary Specification
Equation: UPGRADE_ZSCORE
Date: 05/20/16 Time: 13:32
Success cutoff: C = 0.1203

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	4737	500	5237	0	0	0
P(Dep=1)>C	3917	684	4601	8654	1184	9838
Total	8654	1184	9838	8654	1184	9838
Correct	4737	684	5421	0	1184	1184
% Correct	54.74	57.77	55.10	0.00	100.00	12.03
% Incorrect	45.26	42.23	44.90	100.00	0.00	87.97
Total Gain*	54.74	-42.23	43.07			
Percent Gain...	54.74	NA	48.96			

Appendix D

Presented below are the regression outputs generated when running robustness checks using the probit model. First, the regression output when using the independent variable of interest is presented and is followed by the output for the percentage correctly predicted. Outputs with rating downgrades as dependent variable for both independent variables are presented first, and are followed by the outputs with rating upgrades as dependent variable.

Appendix D.1 – Robustness check with probit model for DD on rating downgrades.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Probit (Quadratic hill climbing)
Date: 05/22/16 Time: 13:15
Sample (adjusted): 3 17400
Included observations: 8869 after adjustments
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.246616	0.018061	-69.02246	0.0000
LAG1DD_WINSORIZE...	-0.059959	0.004875	-12.29966	0.0000

McFadden R-squared	0.023190	Mean dependent var	0.107227
S.D. dependent var	0.309420	S.E. of regression	0.307217
Akaike info criterion	0.666009	Sum squared resid	836.8865
Schwarz criterion	0.667607	Log likelihood	-2951.415
Hannan-Quinn criter.	0.666553	Deviance	5902.830
Restr. deviance	6042.965	Restr. log likelihood	-3021.482
LR statistic	140.1352	Avg. log likelihood	-0.332779
Prob(LR statistic)	0.000000		

Obs with Dep=0	7918	Total obs	8869
Obs with Dep=1	951		

Appendix D.2 – Robustness check with probit model for DD on rating downgrades.

Expectation-Prediction Evaluation for Binary Specification
Equation: PROBIT_DOWN_DD
Date: 05/22/16 Time: 14:31
Success cutoff: C = 0.1072

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	4978	415	5393	0	0	0
P(Dep=1)>C	2940	536	3476	7918	951	8869
Total	7918	951	8869	7918	951	8869
Correct	4978	536	5514	0	951	951
% Correct	62.87	56.36	62.17	0.00	100.00	10.72
% Incorrect	37.13	43.64	37.83	100.00	0.00	89.28
Total Gain*	62.87	-43.64	51.45			
Percent Gain...	62.87	NA	57.63			

Appendix D.3 – Robustness check with probit model for Z-score on rating downgrades.

Dependent Variable: DV_DOWNGRADE
Method: ML - Binary Probit (Quadratic hill climbing)
Date: 05/22/16 Time: 13:16
Sample (adjusted): 3 17400
Included observations: 9838 after adjustments
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.270305	0.017192	-73.88996	0.0000
LAG1ZSCORE_WINSORIZE...	-0.086948	0.015110	-5.754417	0.0000

McFadden R-squared	0.004504	Mean dependent var	0.103476
S.D. dependent var	0.304595	S.E. of regression	0.304237
Akaike info criterion	0.662720	Sum squared resid	910.4201
Schwarz criterion	0.664182	Log likelihood	-3257.919
Hannan-Quinn criter.	0.663215	Deviance	6515.837
Restr. deviance	6545.315	Restr. log likelihood	-3272.658
LR statistic	29.47773	Avg. log likelihood	-0.331157
Prob(LR statistic)	0.000000		

Obs with Dep=0	8820	Total obs	9838
Obs with Dep=1	1018		

Appendix D.4 – Robustness check with probit model for Z-score on rating downgrades.

Expectation-Prediction Evaluation for Binary Specification
Equation: PROBIT_DOWN_ZSCORE
Date: 05/22/16 Time: 14:32
Success cutoff: C = 0.10347

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	5833	506	6339	0	0	0
P(Dep=1)>C	2987	512	3499	8820	1018	9838
Total	8820	1018	9838	8820	1018	9838
Correct	5833	512	6345	0	1018	1018
% Correct	66.13	50.29	64.49	0.00	100.00	10.35
% Incorrect	33.87	49.71	35.51	100.00	0.00	89.65
Total Gain*	66.13	-49.71	54.15			
Percent Gain...	66.13	NA	60.40			

Appendix D.5 – Robustness check with probit model for DD on rating upgrades.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Probit (Quadratic hill climbing)
Date: 05/22/16 Time: 13:17
Sample (adjusted): 3 17400
Included observations: 8869 after adjustments
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.196717	0.017665	-67.74606	0.0000
LAG1DD_WINSORIZE...	0.041587	0.004776	8.707442	0.0000

McFadden R-squared	0.009926	Mean dependent var	0.121209
S.D. dependent var	0.326388	S.E. of regression	0.325477
Akaike info criterion	0.731772	Sum squared resid	939.3281
Schwarz criterion	0.733371	Log likelihood	-3243.042
Hannan-Quinn criter.	0.732316	Deviance	6486.083
Restr. deviance	6551.111	Restr. log likelihood	-3275.555
LR statistic	65.02775	Avg. log likelihood	-0.365660
Prob(LR statistic)	0.000000		
Obs with Dep=0	7794	Total obs	8869
Obs with Dep=1	1075		

Appendix D.6 – Robustness check with probit model for DD on rating upgrades.

Expectation-Prediction Evaluation for Binary Specification
Equation: PROBIT_UP_DD
Date: 05/22/16 Time: 14:33
Success cutoff: C = 0.1212

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	4015	434	4449	0	0	0
P(Dep=1)>C	3779	641	4420	7794	1075	8869
Total	7794	1075	8869	7794	1075	8869
Correct	4015	641	4656	0	1075	1075
% Correct	51.51	59.63	52.50	0.00	100.00	12.12
% Incorrect	48.49	40.37	47.50	100.00	0.00	87.88
Total Gain*	51.51	-40.37	40.38			
Percent Gain...	51.51	NA	45.95			

Appendix D.7 – Robustness check with probit model for Z-score on rating upgrades.

Dependent Variable: DV_UPGRADE
Method: ML - Binary Probit (Quadratic hill climbing)
Date: 05/22/16 Time: 13:18
Sample (adjusted): 3 17400
Included observations: 9838 after adjustments
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-1.176620	0.016476	-71.41442	0.0000
LAG1ZSCORE_WINSORIZE...	0.107554	0.018285	5.882069	0.0000

McFadden R-squared	0.006071	Mean dependent var	0.120350
S.D. dependent var	0.325387	S.E. of regression	0.324608
Akaike info criterion	0.731185	Sum squared resid	1036.420
Schwarz criterion	0.732648	Log likelihood	-3594.700
Hannan-Quinn criter.	0.731681	Deviance	7189.400
Restr. deviance	7233.313	Restr. log likelihood	-3616.656
LR statistic	43.91252	Avg. log likelihood	-0.365389
Prob(LR statistic)	0.000000		

Obs with Dep=0	8654	Total obs	9838
Obs with Dep=1	1184		

Appendix D.8 – Robustness check with probit model for Z-score on rating upgrades.

Expectation-Prediction Evaluation for Binary Specification
Equation: PROBIT_UP_ZSCORE
Date: 05/22/16 Time: 14:34
Success cutoff: C = 0.1203

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=...	4577	479	5056	0	0	0
P(Dep=1)>C	4077	705	4782	8654	1184	9838
Total	8654	1184	9838	8654	1184	9838
Correct	4577	705	5282	0	1184	1184
% Correct	52.89	59.54	53.69	0.00	100.00	12.03
% Incorrect	47.11	40.46	46.31	100.00	0.00	87.97
Total Gain*	52.89	-40.46	41.65			
Percent Gain...	52.89	NA	47.35			