# Have credit ratings become more accurate?

Zvika Afik*

Department of Business Administration
Guilford Glazer Faculty of Business and Management
Ben-Gurion University of the Negev, Israel
Email: afikzv@som.bgu.ac.il

Nitzan Bouhnick

Department of Economics
Ben-Gurion University of the Negev, Israel
Email: nitzan.bouhnick@gmail.com

and

Koresh Galil

Department of Economics
Ben-Gurion University of the Negev, Israel
Email: galilk@exchange.bgu.ac.il

DRAFT: January 30, 2016

*JEL classifications:* G17; G33; G13

*Keywords:* Credit risk; Credit rating; Default prediction

* Corresponding author

# Have credit ratings become more accurate?


## Abstract

Rating agencies play a central role in bond markets. However, the quality of credit ratings is continuously debated. We suggest that the tightening of rating standards observed in prior literature may be partially explained by rating accuracy improvement. We reconfirm that corporate credit ratings have become more stringent over time. Firms with similar accounting profile have a lower rating then they used to have previously.  This paper sheds new light on the evolving quality of credit ratings and their underlying trends. Our analysis shows that ratings are now more correlated with market data than before and less correlated with accounting data than before. Furthermore, we find evidence for improving credit rating quality over time in terms of default prediction. We conclude that rating accuracy has improved over time, and hence, some of prior studies' critique on rating agencies seems outdated.

# 1. Introduction

Previous studies show that corporate credit ratings have become more stringent over time. Firms with similar accounting profile have a lower rating then they used to have previously. As credit ratings are ordinal measures, systematically assigning lower grades to all firms might not be informative nor would it add value to investors.

We conjecture that the change in rating criteria is not merely due to the adoption of more stringent rating rules but it reflects the attempts to provide more accurate credit risk assessments. We illustrate the causality between rating accuracy and the apparent rating standard tightening in a simplifying example. Suppose there are two equally sized groups of firms, high quality (HQ) firms and low-quality (LQ) firms. The firms differ in their probability of default (PD): 5% for HQ firms and 10% for LQ firm. A rating agency employs a rating model that rates firm to A and B for HQ and LQ firms respectively. Suppose the current model is inaccurate and falsely grades a tenth of each group HQ and LQ to the other group, i.e. the wrong rating. Then the average PD would be 5.5% and 9.5% for A and B rated firms respectively. When the rating agency improves its model, supposedly to a perfect model, then the average PD would become 5% and 10% for A and B rated firms respectively. This drop in PD of A-rated firms might be interpreted as standard tightening while in fact it is only a result of the rating model accuracy improvement.

The objective of this work is to test and analyze the quality trends in corporate credit ratings. In contrast to previous studies, we wish to examine whether corporate ratings have improved, and containing now other and more relevant information than before, and whether ratings have become more accurate in their credit risk assessment over time. Specifically, this work addresses the following questions: (i) Does the rating decline and standard tightening trend continue in the 21st century (ii) Is there a negative trend in the PD of high-rated firms; (iii) Do rating agencies change their standards and are ratings now more correlated with market data than before; (iv) Have credit rating quality improved over time in a way that it has a better performance in distinguishing between firms with low PD and high PD?

Our main findings to the above questions are: (i) rating decline continues in the 21$^{st}$ century ; (ii) we find evidence to credit rating changing standards, becoming more stringent over time, in a way that high-rated firms have now lower PD ; (iii) this paper finds that ratings become more correlated to market variables over time (iv) credit rating quality improves over time as evident by their default predictive power.

We believe this paper is the first to show that ratings rely more heavily on non-accounting data and that ratings have become more accurate over time. Furthermore, to maintain compatibility and comparability, in the process of this work we attempt to follow and repeat prior literature methodology. Nevertheless, we also introduce certain technical novelty to this strand of literature, mainly the use of distance-to-default to control for default probability and the area under the receiver operating characteristic curve, on an annual basis, for the assessment of credit rating predictive power.

The rest of the paper proceeds as follows: Section 2 describes the background, Section 3 presents the data and Section 4 describes the methodology. Results are presented and discussed in Section 5 followed by Section 6 which concludes.

## 2. Background

The rating agencies describe how they utilize publically available information and private information revealed to them by the rated firms in their process of rating a specific debt instrument, see for example Standard & Poor's (2008) and S&P's website. This is not a trivial process for the rated firm as it is critical for its debt funding opportunities and affects its capital structure and cost of capital. Faulkender and Petersen (2006) show that, ceteris paribus, a firm with access to the bond market significantly increases its debt compared to similar firms without such access to bond markets. White (2013) provides a broad overview of the rating agencies and their role in financial markets, including historical perspective from their early start to the recent years.

A credit rating agency may apply to the U.S. Securities and Exchange Commission (SEC) for registration as a nationally recognized statistical rating organization (NRSRO). NRSROs' number has fluctuated over the years from seven in the early 1980s down to three during the 1990s and up to 10 NRSROs presently.[1] Beaver et al. (2006) examine whether the properties of bond ratings from NRSROs differ from those of non-certified bond-rating agencies. While NRSROs ratings are used by a variety of constituents, often for regulatory and contractual purposes, ratings from non-certified agencies are used solely for investment advice. Beaver et al. find that NRSROs are generally more conservative than non-certified agencies. Furthermore, Opp et al. (2013) show that introducing rating-contingent regulation that favors highly rated securities may increase or decrease rating informativeness, but unambiguously increases the volume of highly rated securities.

[1] As of September 2015. The list is posted on the SEC's website (http://www.sec.gov/ocr)

Since the 70s, the number of downgrades in corporate bond ratings has exceeded the number of upgrades, leading some to conclude that the credit quality of U.S. corporate debt has declined, see for example Blume, Lim, and Mackinlay (1998, BLM hereafter). Furthermore, they find that the share of S&P AAA rated firms has declined from 8.2% in 1978 to 2.8% in 1995 of all S&P rated firms (our data shows even a lower share of 0.5% in 2010).

BLM apparently were the first to document the declining rating trend phenomenon and to show that it is caused by more stringent rating criteria rather than a response to the evolving accounting data. They provide a structured review of its respective prior literature. Aiming for brevity, below we present prominent examples of later related literature.

BLM study of the rating trend does not address various research questions and is followed by a strand of literature about the rating decline phenomenon. Needless to say that corporate bond rating decline has significant economic consequences. Baghai Servaes and Tamayo (2014, BST hereafter) start by confirming that the rating decline trend continues to 2009 and find empirical evidence that it is mainly caused by rating agencies increasing conservatism. Then they find that "firms affected more by conservatism issue less debt, have lower leverage, hold more cash, are less likely to obtain a debt rating, and experience lower growth." Interestingly, BST also conclude that "firms and capital markets do not perceive the increase in conservatism to be fully warranted."

Among the significant unresolved questions is what causes the trend of corporate bond rating decline in the U.S. A few recent papers partially address this matter, including Jorion et al. (2009), BST, and Alp (2013). A narrower question is whether the rating decline is a result of increasing default probability or it is merely caused by rating agencies tightening standards. BST and Jorion et al. (2009) find no increase in realized default event frequency during the period of the observed credit rating decline. Jorion et al. (2009) find that the downward trend does not apply to speculative-grade issuers and their analysis of investment-grade issuers suggests that it is primarily caused by changes in accounting quality over time. That is, Jorion et al. (2009) conclude that the cause is declining accounting quality and not tightening of the rating agencies standards, hence, contradicting BLM conclusion. On the other hand, Using proprietary longitudinal data, Givoly et al. (2013) examine the change over time in the information content of accounting numbers from the perspective of bondholders. They find that, in contrast to the decline in the information content of accounting numbers to equity holders over time, the information content to bondholders has held steady or risen and suggest that it is related to the increase in reporting conservatism over the last four decades.

Alp (2013) finds a divergent pattern between investment-grade and speculative-grade rating standards from 1985 to 2002 as investment-grade standards tighten and speculative-grade loosen. Later, the analysis shows a structural break toward "more stringent" standards around 2002 for both investment-grade and speculative-grade ratings. Alp (2013) explains this structural break by rating agencies employing more conservative rating practices following Sarbanes-Oxley Act (SOX), responding to the high-profile corporate scandals such as Enron.

# 3. Data

### *3.1 The sample*

The initial sample for this study includes all firms in the merged CRSP-COMPUSTAT database of the period 1986 to 2010 and long-term rating from COMPUSTAT. This period of 25 years is longer than used in prior literature and allows for sufficient overlap with BLM data, which ends in 1995. We include only firms rated by S&P with rating B and higher. With this respect, we differ from BLM that only includes investment-graded firms but unlike Jorion et al. (2009), Alp (2013), and BST (2014) we exclude firms with rating CCC and lower. The speculative ratings have become more prevalent over time in the bond market and such wider spectrum of rating is used in contemporary research (e.g. Mählmann, 2011). After omitting financial firms (SIC 6000-6999) as in Alp (2013) our database includes 15,694 annual observations of 1,534 firms.

### *3.2 The variables*

We follow Jorion et al. (2009) and Alp (2013) in using long-term S&P ratings from Compustat. To facilitate the analysis we assign ascending numbers to the ratings, one to B, two to BB, and so forth ending with six to AAA. Similar to prior research (such as BST), to accommodate the information lag of accounting data we shift the rating data by six months. i.e., the year-end accounting data observation is matched with the rating of six months later. Certain prior research delineate the observations to investment and speculative grades. This generates small sample statistical inference issues in some of the tests, see for example Jorion et al. (2009), therefore we mostly treat the entire spectrum of ratings (AAA-B) as one whole set.

The explanatory variables are of two types - accounting and market variables, and we maintain similarity to BLM definitions to allow for research continuity and comparison.

Following S&P published methodology and other research such as BLM, Jorion et al. (2009), and Mählmann (2011), all the accounting ratios we use, defined below, are moving averages over a window of three years, to smooth local noise and stabilize their values.[2]

*Interest Coverage Ratio* (ICR), the ratio of operating income to interest expenses. Similar to BLM the admissible range of [0, 1] is allocated to four sub-variables, each for a subset of the admissible range. The sub-variable definition and use is presented in the Methodology section, it allows for non-linearity in the ICR influence on the assigned credit rating. Before calculating the average, any negative ratio is set to zero. After calculating the average, any value above 100 is set to 100, following BLM's assumption that a rate higher than 100 conveys no additional information beyond that ceiling.

*Operating Margin,* the ratio of operating income (before depreciation and other deductions) to sales.

*Long Term Debt Leverage,* the ratio of long-term debt to total assets.

*Total Debt Leverage*, the ratio of debt to total assets.

We now turn to the market variables.

*Market Value,* the equity market value of the firm, in million dollars, adjusted by the U.S. consumer price index (CPI) of December 2011 and then converted to its natural logarithm value.

*Market Model Beta* is the commonly used beta, using the last 100 daily equity returns and CRSP (NYSE-NASDAQ) value weighted index. To coincide with the credit rating data, for each observation beta is based on returns shifted by six months, e.g. for accounting data as of 31 December 2008 we calculate beta over the period 4 February 2009 to 30 June 2009. Following to BLM we normalize each beta observation by the average value of all beta observations for the same year.

*Standard Error*[3] from the market model is calculated too, parallel to beta, with the same timing and same normalization procedure. These two variables are assumed to capture and separate the market risk from the idiosyncratic risk.

*Distance to Default* (DD) is a measure often used in default prediction, following Merton (1974). We adopt the specifications suggested by Afik et al (2016) using equation (1). A

---

[2] The ratios are averaged (not the nominator or denominator separately). The average window ends on the observation date, i.e. only information known at the observation date is used.

[3] BST calls it *idiosyncratic risk*.

higher DD indicates a safer firm, less likely to default in the horizon T which equals one year in our application (and it is often the period of choice for researchers and practitioners).

(1) $$DD = \frac{\ln\left(\frac{A}{D}\right)+[\mu_A-0.5\sigma_A^2]T}{\sigma_A\sqrt{T}}$$

Similar to prior literature, for the default threshold we use D = STD+0.5·LTD, where STD is the short term debt (debt maturing in one year) and LTD is the long term debt. A is the asset value which equals the debt value (D) plus the nominal Market Value.[4] The maximum of last year equity return and the risk-free rate is used as a proxy for the asset returns ($\mu_A$). For the risk free rate we use one-year U.S. T-bill rate. All values for equation (1) lag six month after the year-end of the observation, to coincide with other market observables and the credit rating.

Similar to Francis and Schipper (1999) and BST, all explanatory variables, except for ICR, are winsorized at the 99th percentile. Beta and idiosyncratic risk are winsorized prior to normalizing to mitigate the impact of outliers.

Table 1 presents the means of the explanatory variables for each rating class and overall. As can logically be expected, higher ratings, on average, correspond to higher distances to default, higher interest coverages, higher market values and lower debt leverages. All these relations are monotone. Market model beta relation to rating is U-shaped and so is the standard error, though the latter is mostly decreasing as rating increases and it is quite flat for A-AAA ratings. A non-monotone relation between rating and beta is also observed in prior literature (Mählmann, 2011).

[insert TABLE 1 about here]

Table 2 and Figure 1 show the ratings assigned by S&P to U.S. firms on an annual basis in our sample. Panel A presents the number of firms in each rating per year, panel B and Figure 1 show the respective percentage breakdown over time. In 1984 4.2% of our sample are of AAA, 22.9% of AA, and 39.8% are of A rating, these percentages drop significantly over time and are 0.5, 2.5, and 12.3% respectively by 2010. During the same 27 year period, the rating B share has increased from 6.7% to 26.7% while BB and BBB shares increased too, albeit more moderately. Overall, these trends are consistent with BLM, Jorion et al. (2009), Alp (2013) and BST. A possible explanation to these relative changes is the entry of many firms of low

---

[4] Here we do not adjust the market value by the CPI because D is in nominal terms too.

ratings to the sample over time. While this is valid, we can still see a decline in the absolute numbers of firms in the higher rating groups (A, AA, and AAA) over time in panel A of Table 1. Therefore, we need to investigate the changes thoroughly to analyze and understand the sources of the observed trends.

[insert TABLE 2 and FIGURE 1 about here]

# 4. Methodology

## *4.1 Rating versus firm characteristics evolution over time*

We aim to assess whether rating agencies have changed their rating standards over time in two stages. First, we follow the footsteps of prior research in using accounting and market variables to control for firm characteristics over time. Second, we use DD (of equation 1) to control for each firm default probability over time.

### *4.1.1 Controlling for accounting and market variables*

The first stage simply extends prior research analysis into the 21st century, adhering to BLM methodology for comparison, which follows prior research such as Kaplan and Urwitz (1979) and Iskandar-Datta and Emery (1994). While these attend to the selection of optimal credit rating prediction models, BLM focus on the declining U.S. corporate rating, adding yearly dummy variables to the prediction variables. We use *Ordered Probit* (defined in equations 2-3), where the explained variable is the credit rating designated by integers 1, 2, …, 6 for B, BB, .., AAA ratings respectively. We include *Random Effect*s in the regression to address unobserved variables.

(2) $$R_{it} = j\ ,\ \text{when}\ \ \mu_{j-1} < Z_{it} \leq \mu_j$$

where $R_{it}$ is firm i rating in year t, $\mu_j$ ($j = 1,..,5$) is a set of thresholds between rating $j$ and $j+1$, the edge values $\mu_0$ and $\mu_6$ are $-\infty$ and $\infty$ respectively ($\mu_6$ inequality is strict). $Z_{it}$ is a latent variable representing firm i credit risk in year t as follows:

(3) $$Z_{it} = \alpha_t + \beta' X_{it} + \nu_i + \varepsilon_{it}$$

where $\alpha_t$ is the constant term in year t, $\beta$ is a set of slope coefficients, $X_{it}$ is a set of explanatory variables for firm i in year t, $\nu_i$ (the random-effect) are independent and identically distributed $N(0, \sigma_\nu^2)$, and $\varepsilon_{\mathrm{it}}$ is noise. [5]

In previous studies $X_{it}$ mainly includes accounting variables and conclude in decreasing $\alpha_t$ over time, interpreted as the adoption of increasing stringent rating criteria. The accounting variable set of $X_{it}$ that we use includes the operating margin, the long-term debt leverage, the total debt leverage, and the interest coverage ratio (ICR), all are defined in section 2 and the latter requires a special treatment. Similar to BLM, Jorion et al. (2009), Mählmann (2011), and Alp (2013), to allow for non-linearity in the ICR influence, we use four sub-variables defined in equation (4) by the value $ICR_{it}$ of firm *i* in year *t*. This subdivision also allows for better assessment of ICR influence, i.e. when the set of ICR $\beta_j$'s above a certain j are close to zero, it means that only ICR values up to j are significant and an increase above them does not provide additional information about the credit risk of the firm.

(4)

| | $C_{1it}$ | $C_{2it}$ | $C_{3it}$ | $C_{4it}$ |
|---|---|---|---|---|
| $ICR_{it} \in [0,5)$ | $ICR_{it}$ | 0 | 0 | 0 |
| $ICR_{it} \in [5,10)$ | 5 | $ICR_{it} - 5$ | 0 | 0 |
| $ICR_{it} \in [10,20)$ | 5 | 5 | $ICR_{it} - 10$ | 0 |
| $ICR_{it} \in [20,100]$ | 5 | 5 | 10 | $ICR_{it} - 20$ |

The market variable set of $X_{it}$ that we use includes the market value, the beta, and standard error of the market model, all are defined in section 2.

We expect results similar to those of prior studies, where higher credit rating is correlated with higher operating margin, lower long-term debt leverage and higher ICR (though $C_4$ could exhibit an opposite direction due to its insignificant information value), higher market value, lower market model beta, and lower market model standard error. We expect a negative correlation between total debt leverage and the credit rating, yet it is plausible that the result would be of opposite direction due to the high correlation between total debt leverage and the long-term debt leverage. Finally, we expect to see a downward trend in the yearly dummy variables, at least in the period overlapping prior study sample.

[5] BLM control for heteroscedasticity by assuming an error term which is a function of firm size. Alp (2013) shows that the estimation results are robust to heteroscedasticity. For simplicity we omit this modification.

### *4.1.2 Controlling for default probability*

Here we address the question whether the rating decline trend also reflects decreasing default probability over time among high-rated firms. Jorion et al. (2009) claim that the apparently tightening rating standard is only a response to the evolving meaning of accounting data. This subsection aims to examine whether the downward rating trend is real and reflects real change in probabilities of default of the various rating classes. We use the DD (defined in section 3) as a measure for the default probability.

To the best of our knowledge this is the first paper on declining rating that uses DD as a proxy for default probability in its analysis. Prior research of declining rating (such as Jorion et al 2009 and BST) use realized rates for this purpose. We find DD more appropriate for this purpose because: (i) realized default rates provide an average proxy for default risk for the entire sample per year but not for each individual observation of firm-year; (ii) realized default rate is an ex-post measure whereas DD is an ex-ante measure which parallels that of a rating agency; (3) DD provides a wide spectrum of default probability variation while default rates only provide limited number of observations on realized default events.

To achieve our goal we use here the methodology of Ordered Probit with Random Effects of equation (2), where the explained variable is the credit rating and the explaining variables are DD and yearly dummies, see equation (5) which replaces here the definition of $Z_{\text{it}}$ in equation (3).

$$(5) \qquad Z_{\text{it}} = \alpha_t + \beta \cdot DD_{\text{it}} + \nu_i + \varepsilon_{\text{it}}$$

where $DD_{\text{it}}$ is the distance to default of firm $i$ in year $t$. An increase in DD is expected to be positively correlated with the credit rating.

We expect to estimate a positive $\beta$ which supports a higher rating for a higher DD observation. The estimated constants $\alpha_t$ enable us to uncover whether the observed rating trend reflects changing default probability. Relatively stable $\alpha_t$ constants over time indicate that observed rating trends might be related to changing default probabilities. Alternatively, a declining (rising) trend in $\alpha_t$ can be attributed to a change in rating agencies evaluation criteria, assigning a lower (higher) credit rating at a later time for firms with comparable default probabilities and therefore a tightening (loosening) rating standard

## *4.2 Apparent changes in rating agencies standards*

After we identify the trends in credit ratings and following the assessment whether they reflect changing default probabilities, we now turn to assess whether the agencies rating standards evolve over time. Prior research conclusions are mixed. For example, BLM conclude that the declining rating trend is caused by more stringent rating criteria of the agencies. Jorion et al. (2009) on the other hand explain the rating decline by declining quality of accounting data provided by the rated firms. They justify this finding by diminishing explanatory power of accounting data for the ratings assigned to investment grade firms.

We assess these two contradictory explanations based the model of equations (2), (3) yet without the yearly dummy variables. Instead we divide the 25 years of data to five exclusive period, of five years each: 1986-1990, 1991-1995, 1996-2000, 2001-2005, and 2006-2010. This division aims at creating sufficiently large samples and to smooth out short economic fluctuations. We look at two type of results which we describe in the two subsections below.

#### *4.2.1 Assessing the tightening of rating rules*

We are interested to see whether the threshold values $\{\mu_1, \mu_2, \mu_3, \mu_4, \mu_5\}$ change over time and whether a trend can be discerned. Rating standards tightening would be manifested by increasing threshold values over time, as firms would need to exhibit better accounting ratios to be granted the same rating assigned to them earlier. If rating criteria are not becoming tighter over time, then we expect the threshold values to maintain a steady level approximately. This regression also allows us to observe variations in standard changes in each rating over time. For example, if speculative ratings have loosened up to 2002 as suggest by Alp (2013), then we should observe declining $\mu_4$ and $\mu_5$ in that period.

#### *4.2.2 Assessing changes in reliance on (non)accounting information over time*

We divide the variables to two sets: accounting variables and market variables. We aim to analyze trends in the influence of these two sets on rating determination over time. We employ the same method with the same sub-periods of five years each. We repeat the Ordered Probit analysis three times, each round we replace the explanatory variable set, always keeping the assigned ratings as explained variables. First we use both accounting and market variables, then we use only accounting variables, and lastly only market variables. To enable meaningful analysis of the estimated coefficients and reduces estimation biases we omit the long-term debt leverage  from the accounting variable set due to its high correlation with total debt leverage.

The results analysis in this subsection is focused on the explanatory power of the various explanatory variable sets and on the changes in the estimated coefficients. Increased reliance of the rating agencies on market (or accounting) variables would be manifested by an increase in their explanatory power and an increase in the absolute value of their estimated coefficients.

### *4.3 Have credit rating quality improved over time?*

The quality of credit ratings is hard to measure. We attempt to estimate it by assessing its ability to forecast defaults. We assign numbers to credit ratings: 1 to AAA, 2 to AA+,… 21 to C. These ordinal scores can be regarded as proxies to default probabilities. Using a database of actual defaults during one-year following a rating observation date, we assess the predictive power of the ratings.

Threshold ratings may be used by investors, lenders, or regulators to classify firms to high-risk or low-risk categories. The classification might be inaccurate. A false positive (FP) error relates to a solvent firm classified to the high-risk category, whereas a false negative (FN) error relates to a defaulting firm classified to the low-risk category. These are often referred to, by statisticians, as type I and type II errors, and are often estimated by empirical data of false positive and true positive rates (FPR and TPR respectively), for each threshold rating, using a database of assigned rating $score$ observations and their related default/solvency realizations. Consider a threshold value α. TPR, also called hit rate, is the number of defaulting firms classified as high-risk ($score \geq \alpha$) divided by the total number of defaulting firms. FPR, also called false alarm rate, is the number of non-defaulting firms classified as high-risk ($score \geq \alpha$) divided by the total number of solvent firms. There is an obvious tradeoff between these two rates. As one lowers the threshold value, he gains in hit rate (TPR) at the cost of higher false alarm rate (FPR).[6]

The Receiver Operating Characteristic (ROC) curve, a graph of TPR versus FPR, is a tool for comparing powers of alternative default models. Figure 2 shows ROC curves demonstrating the tradeoff between hit rates and false alarm rates for all possible critical values. A random model (with no predictive power) is simply the 45 degrees line. Model A is superior to model B when the ROC curve of A is always above the ROC curve of B. When the curves cross, one may compare the Area Under the Curve (AUC) relative to the alternative models. An AUC value is in the range [0, 1] and the AUC of a random model equals 0.5.

We use annual data to calculate AUCs and regard each AUC as a figure of merit, a measure of the credit rating quality for that year. We then look for a trend, whether the rating

---

[6] Two additional terms that are often used are *sensitivity* for hit rate and 1-*specificity* for false alarm rate.

quality improves or deteriorates over time, or simply seems to fluctuate without an obvious trend.

[insert FIGURE 2 about here]

## 5. Empirical Results

Below we address the results, analyzing each of the questions raised above. We first examine whether rating decline trend continues even after controlling for each firm characteristics and its default probability. Then we explore whether rating standards have changed and whether they rely more on non-accounting data. Then we assess whether rating quality has improved over time.

### *5.1 Rating versus firm characteristics evolution over time*

The preliminary question is whether the declining rating trend, documented in earlier research, continues into the 21st century and the short answer is yes. The results of the ordered Probit model, using accounting and market data in equations (2) and (3), are presented in Table 3 and Figure 3. Most interesting, we find a downward trend in the yearly dummy variables until the financial crisis period. The yearly dummy then fluctuates and seems to continue the downward trend after 2008. These results are similar to BLM's dummies that represent the period 1978-1995 in the overlapping period of the two samples (and also to BST and others). Furthermore, our results show that the downward trend continues into the 21st century.

We now examine the explanatory variables estimated coefficients, which are similar to those of BLM and most of their following research. Conforming to our expectations, higher credit rating is correlated with higher operating margin, lower long-term debt leverage, higher market value, lower market model beta, and lower market model standard error. We find that generally higher credit rating is correlated with higher ICR and statistically significant. Yet, consistent with BLM, Jorion et al. (2009) and Alp (2013) when ICR is very high its marginal effect is slightly negative. $\kappa_4$ ($C_4$ coefficient) is small and statistically insignificant, presumably due to its insignificant information value. $\kappa_2$ ($C_2$ coefficient) is positive, yet small and statistically insignificant too. Though we expect a negative correlation between debt leverage and the credit rating, Table 3 presents a positive (small and statistically insignificant) coefficient of total debt leverage, apparently due to the high correlation between total debt leverage and the long-term debt leverage.

[insert TABLE 3 and FIGURE 3 about here]

Next we assess whether the rating decline trend reflects increasing default probability. We address this question using Ordered Probit of equation (2), where the explained variable is the credit rating and the explaining variables are DD and yearly dummies, see equation (5). Table 4 presents the estimation results and shows that indeed DD's coefficient is positive and statistically significant. The higher is the distance to default of Merton model, the higher is the rating. This is of no surprise. Our focus however, is on the yearly dummy variables which are listed in Table 4 and presented graphically in Figure 4. The figure shows a clear downward trend in the yearly dummies over the sample period of 25 years. Figure 4 is similar to Figure 3, yet not identical and exhibits three positive slopes along the general downward pattern, of which one is at the beginning of the financial crisis.

This result confirms that the trend in rating model is real and reflects changes in the credit quality of firms within each rating class. In this regard, our results conform to the conclusion of BLM, BST, and Alp (2013) and not to those of Jorion et al. (2009). We believe that by using forward-looking market-based measure of default risk (DD), we better capture the default probability of each observation. Therefore, we are able to resolve one of the major concerns in this literature. The quality trend is real and cannot be attributed mainly to evolving meaning of accounting variables. High rated firms have now lower average probabilities of default than

[insert TABLE 4 and FIGURE 4 about here]

Our analysis results thus far show that the trend of declining rating exists and continues into the 21st century after controlling for accounting and market variables and reflects declining default probabilities of default for high-rated firms. We now turn to analyze how the rating agencies have changed their standards over time, which variables prominence increased (decreased) if at all, and whether we can detect changes in the rating accuracy over time.

### *5.2 Assessing changes in rating agencies' standards*

We first assess the tightening rating rules and then study changes in the reliance on accounting and market variables.

#### *5.2.1 Assessing the tightening of rating rules*

We assess the changing standards by estimating $\mu_j$ $(j = 1,..,5)$ - the set of thresholds between rating $j$ and $j + 1$ in equation (2) for the latent variable $Z_{it}$ defined in equation (3). As explained and described in section 3.2 we divide the 25 year period to five sub-periods of five years each and estimate the thresholds for each sub-period. The lower boundary versus time for each rating group is listed in columns in Table 5 and lines in Figure 5. The results show a clear increase of the rating thresholds over time, except during the sub-period 2001-2005 when it flattens, especially for the lower ratings B and BB. For B rated bonds, it even decreases slightly during that sub-period.

[insert TABLE 5 and FIGURE 5 about here]

The observed trend of increasing threshold values over time is an evidence for the rating agencies tightening their rating standards across all rating categories. At later times, a firm needs to exhibit better accounting ratios to be granted the same rating it would have been assigned earlier. The increasing boundaries of the lower ratings ($\mu_1$ and $\mu_2$), paralleling the patterns of investment grade rating boundaries, shows that the tightening standard trend of investment grade firms is observed also in speculative-grade firms. This result differs from that of Jorion et al. (2009).[7] Though Alp (2013) concludes in loosening standards among speculative-grade companies, BST shows that this behavior disappears after controlling for unobserved heterogeneity using fixed effects.

### *5.2.2 Assessing changes in reliance on (non)accounting information over time*

Keeping the assigned ratings as explained variables, we repeat the Ordered Probit analysis three times. In each round we replace the explanatory variable set, using accounting variables, market variables, and both sets together. To analyze changes over time in the influence of the variable sets on rating determination we use the five sub-periods of five years each, defined above. We use the pseudo R-squared of each model (accounting data model, market data model, and accounting and market data model) as a measure for the model explanatory power. The results are presented in Table 6 and Figure 6. Overall, the Pseudo R-squared are compatible to those presented by BST for their ordered Logit models (0.235 and 0.265 in BST). The accounting variable set maintain close to steady R-squared over time, whereas the market variable set results in a monotonically increasing R-squared, except for the period 2001-2005 where it slightly decreases. Combining the two data sets together markedly increases R-squared in all periods with a pattern that generally follows that of the market variable set, albeit at higher values.

[7] We believe that Jorion et al. (2009) results are different because they estimate separately, different models (coefficients) for speculative and investment grades, keeping the model coefficients constant over time (except for the intercepts) while we allow the model to evolve over time (changing from one five-year period to another) yet maintaining the same model coefficients for the entire sample, of all grades, during each five year period.

[insert TABLE 6 and FIGURE 6 about here]

The results show that rating agencies seem to rely more heavily on non-accounting data over time while not ignoring the accounting data whose influence remain relatively steady. Furthermore, in our sample period, the most drastic change has happened during the years 1991-2000 where market variable set R-squared increased steeply and that of accounting variable set decreased slightly.

Our next step is to assess the changing weight of each variable on the observed ratings using accounting and market data for the explanatory variables in an ordered Probit model, explaining ratings in five sub-periods. The estimated coefficients (and their t-statistics in parentheses) are listed in Table 7 for each of the sub-periods. The explanatory variables are those that are used for Table 3 above, except for the interest coverage ratio that is used here as is, without delineation to four sub-variables. Figures 7 and 8 are graphical representations of certain lines of Table 7. Figure 7 shows that the coefficients of total debt leverage and operating margin decline in their absolute values over time. Figure 8, on the other hand, shows that market variables' coefficients absolute values increase over time, most prominent is the increase of the market model standard error which is often referred to as the idiosyncratic risk of the firm's equity. To our knowledge this paper is the first to show such evolution of explanatory variables over time.

[insert TABLE 7 and FIGUREs 7,8 about here]

The above analysis results show that rating standards and methodology evolve over time. That apparently, more (less) emphasis is put on market (accounting) data recently compared to the past and that rating assignment criteria have become more stringent over time. The remaining open question for this work is whether those changes have affected the quality of the assigned credit ratings.

### *5.3 Have credit rating quality improved over time?*

Assessing credit rating quality is a fuzzy task with a variety of connotations. As presented in the methodology chapter, we focus on the ability of credit ratings to predict default events in the following year. We use the AUC of the ROC curve as a figure of merit for the predictive power of the model. While the rating agencies model is latent, its outcomes, the credit scores are public information and so are the default realizations. Table 8 and Figure 9 show the evolution of the AUC over time.

For the analysis of this subsection we need two sets of data: annual rating observations and default events during each of the following years. This sample includes 52,776 observations from 1985-2013 of all firms in the CRSP-Compustat merged database with S&P ratings. The overall number of defaults is 1,123 during the years 1986-2014. Each observation includes the S&P rating at the end of the calendar year and an indication whether the default defaulted within the subsequent year. Here we use the fine-grained rating information including the +/- rating notches.

The table and figure show the AUC evolution over time and the respective 90% confidence intervals. Figure 9 shows also the AUC estimated linear trend which is positive, indicating a generally improving rating accuracy trend over the long period of almost three decades.

[insert TABLE 8 and FIGURE 9 about here]

## 6. Summary and conclusions

The above results show that the downward credit rating trend continues into the 21st century even after controlling for accounting and market variables. That is, firms with identical accounting and market variables are rated lower at a later time then they would have been rated earlier. This finding reaffirms prior literature results in our extended database. We then add to prior research a new test in which we replace the accounting and market variables with firms' default risk measured by Merton model. An ordered Probit model reveals a downward trend of credit rating even after controlling for the default risk. Thus, more recently a firm is assigned a lower rating than earlier when the default risk is held constant. This result confirms that the change in rating standard is real and does not only reflect an evolving meaning of accounting variables.

These tests show the annual decline in the assigned rating for the entire sample. Our next step is to zoom-in and focus on the apparent rating rules that are revealed by the data. Using the latent score of accounting and market variables in the ordered Probit model, we estimate the lower-bound score for each rating in our sample. This is done five times for each period of five years in our 25-year sample. The estimated boundaries increase over time, showing that the rating assignment rules of the rating agencies become increasingly more stringent in all the rating groups during our 25-year sample. During the period of 2001-2005 this trends weakens for all ratings, and for the lower rating even reverses direction. After this period and for the remaining years in the sample, the rating assignment boundaries continue to increase. Such a change during 2001-2005 could be interpreted, for example, as a change in rating assignment standards, especially for the lower rating group, or to be explained by changes in accounting standards (i.e. becoming more conservative). Possibly this change is related to

Sarbanes-Oxley Act of 2002 (SOX) and changes in the financial market, firms, and public sentiments during and following the accounting scandals of the early 2000th.[8] Furthermore, except for this specific period we find evidence for credit rating assignments becoming more stringent over time in all periods and for all rating groups, including speculative ratings. The latter is different than some of Jorion et al (2009) results, which find that the rating assignments rules do not change much for speculative graded bonds.

Our next analysis identifies a trend of ratings' increased reliance on market data and decreasing correlation with accounting variables. It is a new contribution of this research to the literature. Another new contribution, to the best of our knowledge, is of higher importance to our view – we find that in our 25-year sample, ratings become more accurate in predicting defaults over time. Thus, the above changes are not simply shifting standards on an arbitrary scale, they improve the signaling quality of the ratings assigned by the agencies.

This paper shows that the downward credit rating trend continues into the 21st century after controlling for accounting and market variables and even after controlling for each firm default probability. These findings are an extension and robustness test relative to earlier research. The paper shows clearly that tightening of rating rules is a general trend applicable to all rating groups, B to AAA. We do not see significant differences between rating threshold patterns of investment grade and speculative grade bonds over a period of 25 years.

Our analysis shows a trend of ratings' increased reliance on market data and decreasing correlation with accounting variables. This is a new contribution of this research to the literature. Probably a more significant new contribution is that in our 25-year sample, ratings show increasing accuracy in predicting defaults over time. These findings may lead to a conclusion that the above changes are evidence of the agencies efforts to enhance the rating quality, improving their rating accuracy and timeliness. Beyond the above listed contributions we also believe that this paper is the first to use distance-to-default and the area under the ROC curve in this strand of literature.

Overall, all these findings are consistent with our main hypothesis. Rating quality has improved over time by increasing reliance on non-accounting data and perhaps better use of

8 Quoting the U.S. government publication of SOX: "An act. To protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws, and for other purposes." 107th Congress Public Law 204, July 30, 2002, see www.gpo.gov/fdsys/pkg/PLAW-107publ204/html/PLAW-107publ204.htm.

accounting data. The tightening rating standards as observed in previous literature may only be a side effect of the higher accuracy of ratings.

We also conclude that some of the critique based on historical rating performance might be outdated. Credit ratings appear valuable, informational, and improving despite the negative findings of historical studies. Such results are significant for researchers, practitioners, and even the regulators.

## References


Afik Z., O. Arad and K. Galil (2016). Using Merton model for default prediction: an empirical assessment of selected alternatives. *Journal of Empirical Finance* 35, 43-67.

Alp, A. (2013). Structural shifts in credit rating standards. *The Journal of Finance*, 68(6), 2435-2470.

Baghai, R.P., H. Servaes, A. and Tamayo (2014). Have rating agencies become more conservative? Implications for capital structure and debt pricing. *The Journal of Finance*, 69(5), 1961-2005.

Beaver, W.H., C. Shakespeare, and M.T. Soliman (2006). Differential properties in the ratings of certified versus non-certified bond-rating agencies. *Journal of Accounting and Economics,* 42, 303-334.

Bharath, S.T. and T. Shumway (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, 21(3), 1339-1369.

Blume, M.E., F. Lim, and C. Mackinlay (1998). The declining credit quality of U.S. corporate debt: Myth or reality. *Journal of Finance*, 53, 1389-1413.

DeLong, E.R., D.M. DeLong, and D.L. Clarke-Pearson (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44, 837-845.

Faulkender, M. and M.A. Petersen (2006). Does the Source of Capital affect Capital Structure? *Review of Financial Studies*, 19, 45-79.

Francis, J. and K. Schipper (1999). Have financial statements lost their relevance? *Journal of Accounting Research*, 319-352.

Givoly, D., C. Hayn, and S. Katz (2013). The Changing Relevance of Accounting Numbers to Debt Holders over Time. working paper available at SSRN.

Iskandar-Datta, M.E., and D.R. Emery (1994). An empirical investigation of the role of indenture provisions in determining bond ratings. *Journal of Banking and Finance*, 18, 93–111.

Jorion, P., Z. Liu, and C. Shi (2005). Informational effects of regulation FD: evidence from rating agencies. *Journal of Financial Economics*, 76, 309-330.

Jorion, P., C. Shi, and S. Zhang (2009). Tightening credit standards: the role of accounting quality. *Review of Accounting Studies*, 14(1), 123-160.

Kaplan, R. S., and Urwitz, G., 1979. "Statistical models of bond ratings: A methodological inquiry" Journal of Business 52, 231–261.

Mählmann, T. (2011). Is There a Relationship Benefit in Credit Ratings? *Review of Finance* 15, 475-510.

Merton, R. C., 1974."On the Pricing of Corporate Debt: The Risk Structure of Interest Rates" Journal of Finance 29, 449–70.

Opp, C.C., M.M. Opp, and M. Harris (2013)."Rating agencies in the face of regulation. *Journal of Financial Economics*, 108(1), 46-61.

Standard & Poor's, 2008 "Corporate Ratings Criteria".

Standard & Poor's, 2013, About credit ratings,

http://www.standardandpoors.com/aboutcreditratings/RatingsManual_PrintGuide.html, last accessed July 29, 2014.

White, L. J. ,2013. "Credit Rating Agencies: An Overview" Annu. Rev. Financ. Econ. 5, 93-122.

# *Tables and Figures*

**Table 1 – Descriptive Statistics**

This table presents means for the explanatory variables for each rating class and overall. The statistics are calculated using a panel data sample of 15694 observations from 1986 through 2010. The sample includes all firms in the CRSP-Compustat merged database with ratings of B and above, excluding financial firms (SIC codes between 6000 to 6999). Interest Coverage is EBIT divided by interest expenses. Operating Margin is EBITDA divided by Sales. LT Debt Leverage is Long-term Debt divided by Total Assets. Total Debt Leverage is Total Debt divided by Total Assets. Market Value is the natural logarithm of Market Value of Equity, in million dollars, adjusted by the U.S. CPI of December 2011. Market Model Beta is estimated using daily stock return. Standard Error is the standard error of the market model using daily stock returns. DD is Distance to Default of equation (1). All variables are averages over the previous three years, winsorized at the 1% and 99% level. Ratings and market data are dated six months after the date of the annual fiscal reports.

| | **Rating** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Total** | **AAA** | **AA** | **A** | **BBB** | **BB** | **B** |
| Interest Coverage | 7.34 | 25.2 | 14.57 | 10.34 | 7.05 | 5.74 | 3.64 |
| Operating Margin | 0.17 | 0.27 | 0.2 | 0.19 | 0.19 | 0.18 | 0.12 |
| LT Debt Leverage | 0.31 | 0.08 | 0.16 | 0.22 | 0.27 | 0.36 | 0.43 |
| Total Debt Leverage | 0.35 | 0.16 | 0.23 | 0.27 | 0.31 | 0.4 | 0.47 |
| Market Value | 7.41 | 10.09 | 8.92 | 8.44 | 7.95 | 6.88 | 5.82 |
| Market Model Beta | 1.00 | 1.12 | 1.02 | 0.91 | 0.9 | 1.03 | 1.16 |
| Standard Error | 1.00 | 0.78 | 0.75 | 0.74 | 0.82 | 1.07 | 1.47 |
| DD | 8.29 | 15.21 | 13.2 | 11.09 | 8.98 | 6.54 | 5.09 |
| Number of Obs. | 15,694 | 222 | 902 | 3,073 | 4,215 | 3,999 | 3,283 |

## Table 2 – Sample by S&P rating and year

The sample includes 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). Panel A shows the number of firms in each rating class and Panel B shows the percentage breakdown by year.

**Panel A: number of firms in each rating class**

| | Rating | | | | | | |
|---|---|---|---|---|---|---|---|
| **Year** | **AAA** | **AA** | **A** | **BBB** | **BB** | **B** | **Total** |
| 1984 | 12 | 65 | 113 | 52 | 23 | 19 | 284 |
| 1985 | 12 | 63 | 118 | 58 | 30 | 24 | 305 |
| 1986 | 12 | 61 | 119 | 71 | 36 | 27 | 326 |
| 1987 | 11 | 64 | 115 | 79 | 44 | 32 | 345 |
| 1988 | 13 | 59 | 127 | 80 | 45 | 38 | 362 |
| 1989 | 14 | 56 | 129 | 93 | 45 | 45 | 382 |
| 1990 | 14 | 58 | 125 | 105 | 45 | 44 | 391 |
| 1991 | 13 | 58 | 130 | 104 | 58 | 37 | 400 |
| 1992 | 14 | 57 | 135 | 117 | 68 | 39 | 430 |
| 1993 | 13 | 54 | 138 | 125 | 94 | 61 | 485 |
| 1994 | 13 | 52 | 143 | 148 | 118 | 90 | 564 |
| 1995 | 11 | 52 | 159 | 149 | 139 | 95 | 605 |
| 1996 | 10 | 49 | 160 | 175 | 163 | 122 | 679 |
| 1997 | 10 | 42 | 171 | 207 | 178 | 160 | 768 |
| 1998 | 10 | 44 | 168 | 219 | 211 | 211 | 863 |
| 1999 | 10 | 40 | 152 | 237 | 217 | 208 | 864 |
| 2000 | 9 | 28 | 145 | 236 | 218 | 208 | 844 |
| 2001 | 8 | 26 | 125 | 228 | 207 | 205 | 799 |
| 2002 | 7 | 20 | 114 | 218 | 209 | 179 | 747 |
| 2003 | 7 | 17 | 111 | 203 | 223 | 156 | 717 |
| 2004 | 6 | 16 | 109 | 196 | 229 | 167 | 723 |
| 2005 | 5 | 16 | 107 | 191 | 240 | 159 | 718 |
| 2006 | 5 | 14 | 93 | 183 | 238 | 181 | 714 |
| 2007 | 5 | 16 | 91 | 189 | 232 | 183 | 716 |
| 2008 | 6 | 14 | 86 | 184 | 220 | 181 | 691 |
| 2009 | 4 | 14 | 78 | 182 | 191 | 167 | 636 |
| 2010 | 3 | 16 | 78 | 185 | 181 | 169 | 632 |
| **Total** | **262** | **1,106** | **3,481** | **4,549** | **4,262** | **3,478** | **15,990** |

**Panel B: percentage breakdown of firms in each rating class**

| | Rating | | | | | | |
|---|---|---|---|---|---|---|---|
| **Year** | **AAA** | **AA** | **A** | **BBB** | **BB** | **B** | **Total** |
| 1984 | 4.2 | 22.9 | 39.8 | 18.3 | 8.1 | 6.7 | 100 |
| 1985 | 3.9 | 20.7 | 38.7 | 19.0 | 9.8 | 7.9 | 100 |
| 1986 | 3.7 | 18.7 | 36.5 | 21.8 | 11.0 | 8.3 | 100 |
| 1987 | 3.2 | 18.6 | 33.3 | 22.9 | 12.8 | 9.3 | 100 |
| 1988 | 3.6 | 16.3 | 35.1 | 22.1 | 12.4 | 10.5 | 100 |
| 1989 | 3.7 | 14.7 | 33.8 | 24.3 | 11.8 | 11.8 | 100 |
| 1990 | 3.6 | 14.8 | 32.0 | 26.9 | 11.5 | 11.3 | 100 |
| 1991 | 3.3 | 14.5 | 32.5 | 26.0 | 14.5 | 9.3 | 100 |
| 1992 | 3.3 | 13.3 | 31.4 | 27.2 | 15.8 | 9.1 | 100 |
| 1993 | 2.7 | 11.1 | 28.5 | 25.8 | 19.4 | 12.6 | 100 |
| 1994 | 2.3 | 9.2 | 25.4 | 26.2 | 20.9 | 16.0 | 100 |
| 1995 | 1.8 | 8.6 | 26.3 | 24.6 | 23.0 | 15.7 | 100 |
| 1996 | 1.5 | 7.2 | 23.6 | 25.8 | 24.0 | 18.0 | 100 |
| 1997 | 1.3 | 5.5 | 22.3 | 27.0 | 23.2 | 20.8 | 100 |
| 1998 | 1.2 | 5.1 | 19.5 | 25.4 | 24.4 | 24.4 | 100 |
| 1999 | 1.2 | 4.6 | 17.6 | 27.4 | 25.1 | 24.1 | 100 |
| 2000 | 1.1 | 3.3 | 17.2 | 28.0 | 25.8 | 24.6 | 100 |
| 2001 | 1.0 | 3.3 | 15.6 | 28.5 | 25.9 | 25.7 | 100 |
| 2002 | 0.9 | 2.7 | 15.3 | 29.2 | 28.0 | 24.0 | 100 |
| 2003 | 1.0 | 2.4 | 15.5 | 28.3 | 31.1 | 21.8 | 100 |
| 2004 | 0.8 | 2.2 | 15.1 | 27.1 | 31.7 | 23.1 | 100 |
| 2005 | 0.7 | 2.2 | 14.9 | 26.6 | 33.4 | 22.1 | 100 |
| 2006 | 0.7 | 2.0 | 13.0 | 25.6 | 33.3 | 25.4 | 100 |
| 2007 | 0.7 | 2.2 | 12.7 | 26.4 | 32.4 | 25.6 | 100 |
| 2008 | 0.9 | 2.0 | 12.4 | 26.6 | 31.8 | 26.2 | 100 |
| 2009 | 0.6 | 2.2 | 12.3 | 28.6 | 30.0 | 26.3 | 100 |
| 2010 | 0.5 | 2.5 | 12.3 | 29.3 | 28.6 | 26.7 | 100 |

## Table 3 – Ordered Probit model for ratings with accounting and market data

The estimates are for the random-effect ordered Probit model parameters of equations (2) and (3), using a panel data sample of 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). Interest Coverage is EBIT divided by Interest expenses, divided to four sub-variables as defined in equation (4). Operating Margin is EBITDA divided by Sales. LT Debt Leverage is Long-term Debt divided by Total Assets. Total Debt Leverage is also divided by Total Assets. Market Value is the natural logarithm of Market Value of Equity, in million dollars, adjusted by the U.S. CPI of December 2011. Market Model Beta is estimated using daily stock return. Standard Error is the standard error of the market model using daily stock returns. All variables are averages over the previous three years, winsorized at the 1% and 99% level. Ratings and market data are dated six months after the date of the annual fiscal reports. (*) and (**) denote statistical significance at the 5% and 1% levels respectively.

| | Coefficient | Std. Err. | P-value |
|---|---|---|---|
| Interest coverage $\kappa_1$ | 0.254** | 0.012 | 0.000 |
| $\kappa_2$ | 0.001 | 0.011 | 0.946 |
| $\kappa_3$ | 0.031** | 0.007 | 0.000 |
| $\kappa_4$ | -0.003 | 0.002 | 0.210 |
| Operating Margin | 0.764** | 0.133 | 0.000 |
| Total Debt Leverage | 0.121 | 0.287 | 0.673 |
| LT Debt Leverage | -2.456** | 0.290 | 0.000 |
| Market Value | 0.637** | 0.014 | 0.000 |
| Market Model beta | -0.141** | 0.021 | 0.000 |
| Standard error | -0.222** | 0.026 | 0.000 |
| <u>Year Dummies</u> | | | |
| 1986 | 0.000 | | |
| 1987 | -0.049 | 0.091 | 0.588 |
| 1988 | -0.232** | 0.090 | 0.010 |
| 1989 | -0.255** | 0.090 | 0.005 |
| 1990 | -0.355** | 0.090 | 0.000 |
| 1991 | -0.423** | 0.089 | 0.000 |
| 1992 | -0.689** | 0.087 | 0.000 |
| 1993 | -0.868** | 0.085 | 0.000 |
| 1994 | -1.024** | 0.084 | 0.000 |
| 1995 | -1.290** | 0.084 | 0.000 |
| 1996 | -1.489** | 0.083 | 0.000 |
| 1997 | -1.654** | 0.083 | 0.000 |
| 1998 | -1.665** | 0.083 | 0.000 |
| 1999 | -1.751** | 0.085 | 0.000 |
| 2000 | -1.993** | 0.087 | 0.000 |
| 2001 | -2.165** | 0.088 | 0.000 |
| 2002 | -2.231** | 0.088 | 0.000 |
| 2003 | -2.510** | 0.090 | 0.000 |
| 2004 | -2.780** | 0.091 | 0.000 |
| 2005 | -3.108** | 0.093 | 0.000 |
| 2006 | -3.301** | 0.094 | 0.000 |
| 2007 | -3.223** | 0.094 | 0.000 |
| 2008 | -3.143** | 0.094 | 0.000 |
| 2009 | -3.217** | 0.096 | 0.000 |
| 2010 | -3.295** | 0.098 | 0.000 |

## Table 4 – Ordered Probit model for ratings with 'Distance to Default'

The estimates are for the random-effect ordered Probit model parameters of equations (2) and (5), using a panel data sample of 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). DD is Distance to Default of equation (1), averaged on the previous three years, winsorized at the 1% and 99% level. Ratings and market data are dated six months after the date of the annual fiscal reports. (*) and (**) denote statistical significance at the 5% and 1% levels respectively.

| | Coefficient | Std. Err. | P-value |
|---|---|---|---|
| DD | 0.050** | 0.002 | 0.000 |
| Year Dummies | | | |
| 1986 | 0.000 | | |
| 1987 | -0.069 | 0.088 | 0.432 |
| 1988 | -0.204* | 0.087 | 0.019 |
| 1989 | -0.146 | 0.087 | 0.094 |
| 1990 | -0.130 | 0.086 | 0.133 |
| 1991 | -0.180* | 0.085 | 0.035 |
| 1992 | -0.350** | 0.084 | 0.000 |
| 1993 | -0.375** | 0.082 | 0.000 |
| 1994 | -0.488** | 0.081 | 0.000 |
| 1995 | -0.509** | 0.080 | 0.000 |
| 1996 | -0.545** | 0.078 | 0.000 |
| 1997 | -0.560** | 0.077 | 0.000 |
| 1998 | -0.500** | 0.077 | 0.000 |
| 1999 | -0.569** | 0.078 | 0.000 |
| 2000 | -0.747** | 0.079 | 0.000 |
| 2001 | -0.969** | 0.080 | 0.000 |
| 2002 | -1.102** | 0.080 | 0.000 |
| 2003 | -1.216** | 0.081 | 0.000 |
| 2004 | -1.309** | 0.081 | 0.000 |
| 2005 | -1.410** | 0.081 | 0.000 |
| 2006 | -1.476** | 0.082 | 0.000 |
| 2007 | -1.267** | 0.081 | 0.000 |
| 2008 | -1.302** | 0.082 | 0.000 |
| 2009 | -1.381** | 0.083 | 0.000 |
| 2010 | -1.413** | 0.084 | 0.000 |

**Table 5 – Estimates of lower boundaries for rating categories using accounting and market data**

This table shows the estimates for the lower-boundaries for rating category parameters from five ordered Probit models using accounting and market data. Each model is for a different sub-period. The overall sample includes 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). The explanatory variables are Interest coverage, Operating margin, LT debt leverage, Total debt leverage, Market value, Market model beta and Market model standard error.

| Period | *BB* | *BBB* | *A* | *AA* | *AAA* |
|---|---|---|---|---|---|
| 1986-1990 | -0.22 | 0.62 | 1.65 | 3.01 | 4.38 |
| 1991-1995 | 0.59 | 1.7 | 2.77 | 4.1 | 5.34 |
| 1996-2000 | 0.97 | 2.2 | 3.47 | 4.87 | 5.95 |
| 2001-2005 | 0.77 | 2.19 | 3.52 | 5.07 | 5.96 |
| 2006-2010 | 1.57 | 3.05 | 4.54 | 5.93 | 6.90 |

**Table 6 – Pseudo R-squared of accounting and market data models**

This table shows the Pseudo R-squared of three types of three ordered Probit models explaining ratings. Each model differs in its explanatory variables: (1) accounting with market data, (2) accounting data and (3) market data. The overall sample includes 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). The accounting data are Interest coverage, Operating margin and, Total debt leverage. The market variable data are Market value, Market model beta and Market model standard error. The results are reported for five sub-periods.

| Period | Accounting and Market variables | Accounting variables | Market variables |
|---|---|---|---|
| 1986-1990 | 0.2315 | 0.1365 | 0.1380 |
| 1991-1995 | 0.2471 | 0.1061 | 0.1884 |
| 1996-2000 | 0.2975 | 0.1268 | 0.2549 |
| 2001-2005 | 0.2950 | 0.1108 | 0.2496 |
| 2006-2010 | 0.3260 | 0.1205 | 0.2798 |

## Table 7 – Coefficient estimates in sub-periods

This table shows the coefficient estimate of the ordered Probit model explaining ratings in five sub-periods. The numbers in parentheses are the respective t-values. The overall sample includes 15,694 observations from 1986 through 2010 of all firms in the CRSP-Compustat merged database with ratings of B or above, excluding financial firms (SIC code between 6000 to 6999). Interest Coverage is EBIT divided by Interest expenses. Operating Margin is EBITDA divided by Sales. LT Debt is Long-term Debt divided by Total Assets. Total Debt is also divided by Total Assets. Market Value is the natural logarithm of Market Value of Equity, in million dollars, adjusted by the U.S. CPI of December 2011. Market Model Beta is estimated using daily stock return. Standard Error is the standard error of the market model using daily stock returns. All variables are averages over the previous three years, winsorized at the 1% and 99% level. Ratings and market data are dated six months after the date of the annual fiscal reports. (*) and (**) denote statistical significance at the 5% and 1% levels respectively.

| | Period | | | | |
|---|---|---|---|---|---|
| | 1986-1990 | 1991-1995 | 1996-2000 | 2001-2005 | 2006-2010 |
| Interest Coverage | -0.002 | 0.007 | 0.005 | 0.005 | 0.009 |
| | (-0.74) | (3.52)** | (3.13)** | (3.13)** | (5.88)** |
| Operating Margin | 2.607 | 1.303 | 1.178 | 0.581 | 0.28 |
| | (10.73)** | (7.56)** | (9.35)** | (4.15)** | (1.84) |
| Total debt Leverage | -4.376 | -2.221 | -1.789 | -1.942 | -1.884 |
| | (-16.02)** | (-20.42)** | (-15.97)** | (-14.95)** | (-13.53)** |
| Market Value | 0.443 | 0.486 | 0.488 | 0.465 | 0.525 |
| | (20.34)** | (29.02)** | (37.99)** | (32.30)** | (32.28)** |
| Market Model Beta | -0.145 | -0.242 | -0.268 | -0.4 | -0.383 |
| | (-2.66)** | (-7.05)** | (-10.51)** | (-10.36)** | (-7.21)** |
| Standard Error | -0.174 | -0.278 | -0.505 | -0.602 | -0.601 |
| | (-3.57)** | (-6.18)** | (-11.03)** | (-13.84)** | (-11.70)** |
| Number of Obs. | 1,867 | 2,725 | 4,104 | 3,737 | 3,261 |

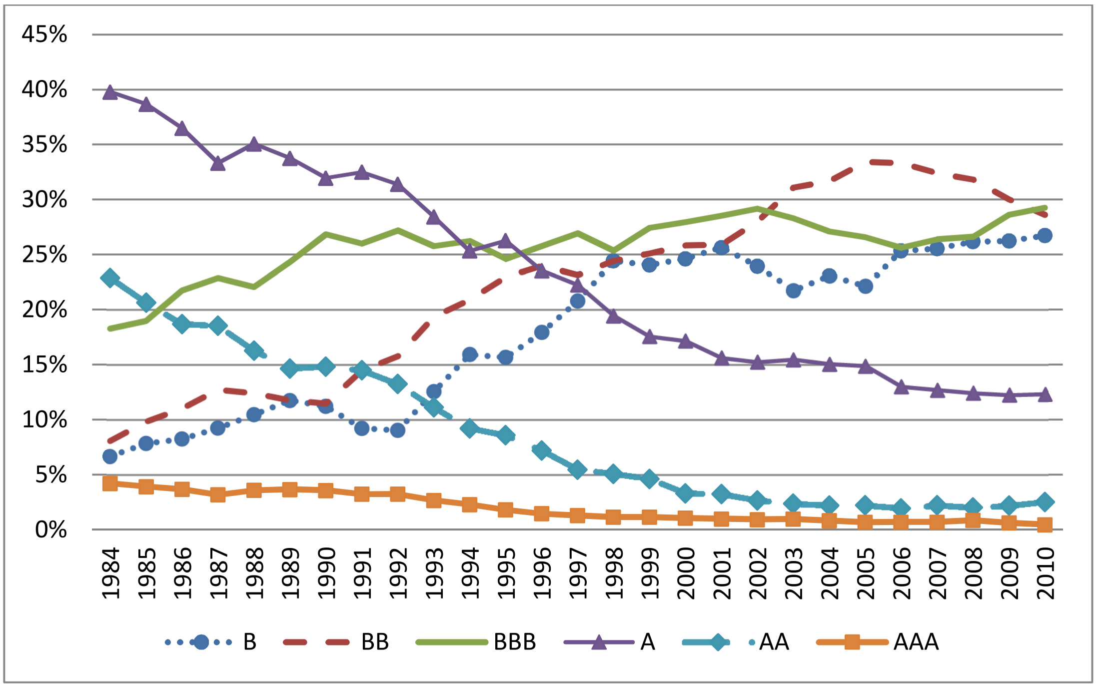
## Table 8 – Area under curve

This table shows the Area Under Curve (AUC) of Receiver Operating Characteristics (ROC) curve for prediction of default using credit ratings. The sample includes 52,776 observations from 1985-2013 of all firms in the CRSP-Compustat merged database with S&P ratings. The overall number of defaults is 1,123 for the years 1986-2014. Each observation includes the S&P rating at the end of the calendar year and an indication whether the default defaulted within the subsequent year.

| Years | Observations | Defaults within 1 years | AUC | AUC Confidence Interval | |
|---|---|---|---|---|---|
| | | | | 5% | 95% |
| 1985 | 1249 | 19 | 0.8356 | 0.7793 | 0.8919 |
| 1986 | 1411 | 16 | 0.8948 | 0.8590 | 0.9305 |
| 1987 | 1467 | 19 | 0.8799 | 0.8268 | 0.9330 |
| 1988 | 1451 | 27 | 0.8690 | 0.8149 | 0.9231 |
| 1989 | 1419 | 46 | 0.8519 | 0.8081 | 0.8957 |
| 1990 | 1327 | 48 | 0.8673 | 0.8313 | 0.9034 |
| 1991 | 1290 | 17 | 0.9328 | 0.9158 | 0.9497 |
| 1992 | 1350 | 12 | 0.8941 | 0.8193 | 0.9689 |
| 1993 | 1464 | 9 | 0.9163 | 0.8608 | 0.9717 |
| 1994 | 1520 | 18 | 0.8735 | 0.7904 | 0.9565 |
| 1995 | 1618 | 12 | 0.9165 | 0.8775 | 0.9555 |
| 1996 | 1759 | 18 | 0.9375 | 0.9059 | 0.9690 |
| 1997 | 1908 | 33 | 0.9161 | 0.8820 | 0.9502 |
| 1998 | 2125 | 59 | 0.9027 | 0.8718 | 0.9336 |
| 1999 | 2236 | 70 | 0.8615 | 0.8273 | 0.8957 |
| 2000 | 2262 | 125 | 0.8607 | 0.8301 | 0.8913 |
| 2001 | 2242 | 95 | 0.8463 | 0.8118 | 0.8808 |
| 2002 | 2193 | 62 | 0.9083 | 0.8835 | 0.9331 |
| 2003 | 2199 | 28 | 0.9302 | 0.8991 | 0.9614 |
| 2004 | 2214 | 23 | 0.8922 | 0.8321 | 0.9523 |
| 2005 | 2164 | 13 | 0.9437 | 0.8981 | 0.9893 |
| 2006 | 2130 | 10 | 0.9492 | 0.8768 | 1.0000 |
| 2007 | 2045 | 54 | 0.8689 | 0.8130 | 0.9247 |
| 2008 | 1975 | 129 | 0.9114 | 0.8828 | 0.9400 |
| 2009 | 1939 | 53 | 0.9520 | 0.9058 | 0.9981 |
| 2010 | 1950 | 34 | 0.9480 | 0.9024 | 0.9936 |
| 2011 | 1957 | 35 | 0.9770 | 0.9601 | 0.9939 |
| 2012 | 1940 | 24 | 0.9573 | 0.8888 | 1.0000 |
| 2013 | 1972 | 15 | 0.9836 | 0.9760 | 0.9913 |

## Figure 1 – Sample by S&P rating and year

This figure shows the percentage of firms in each S&P rating class in the period 1984-2010




## Figure 2: Illustration of ROC

Demonstrating curves of true positive rate (TPR) versus false positive rate (FPR) for two models and for a random order.

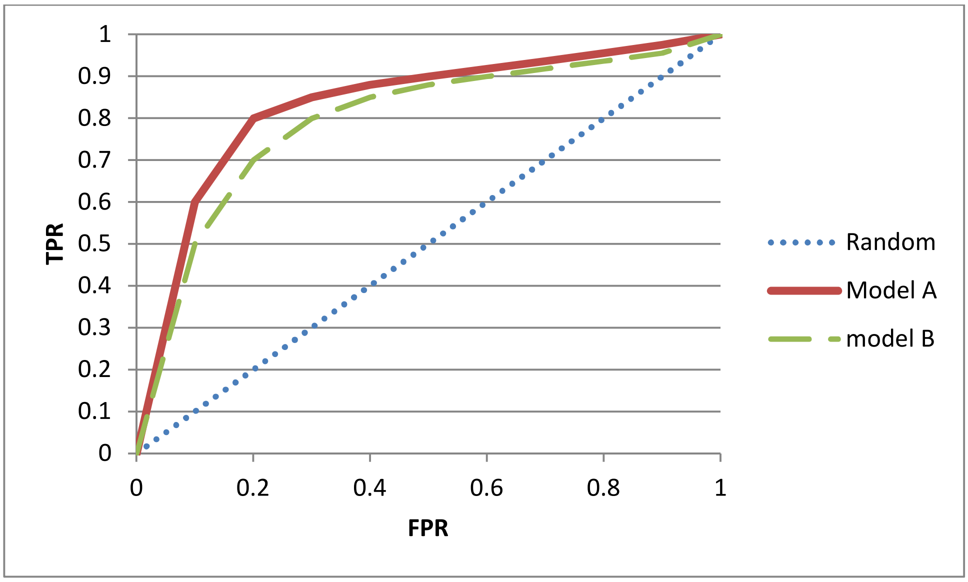
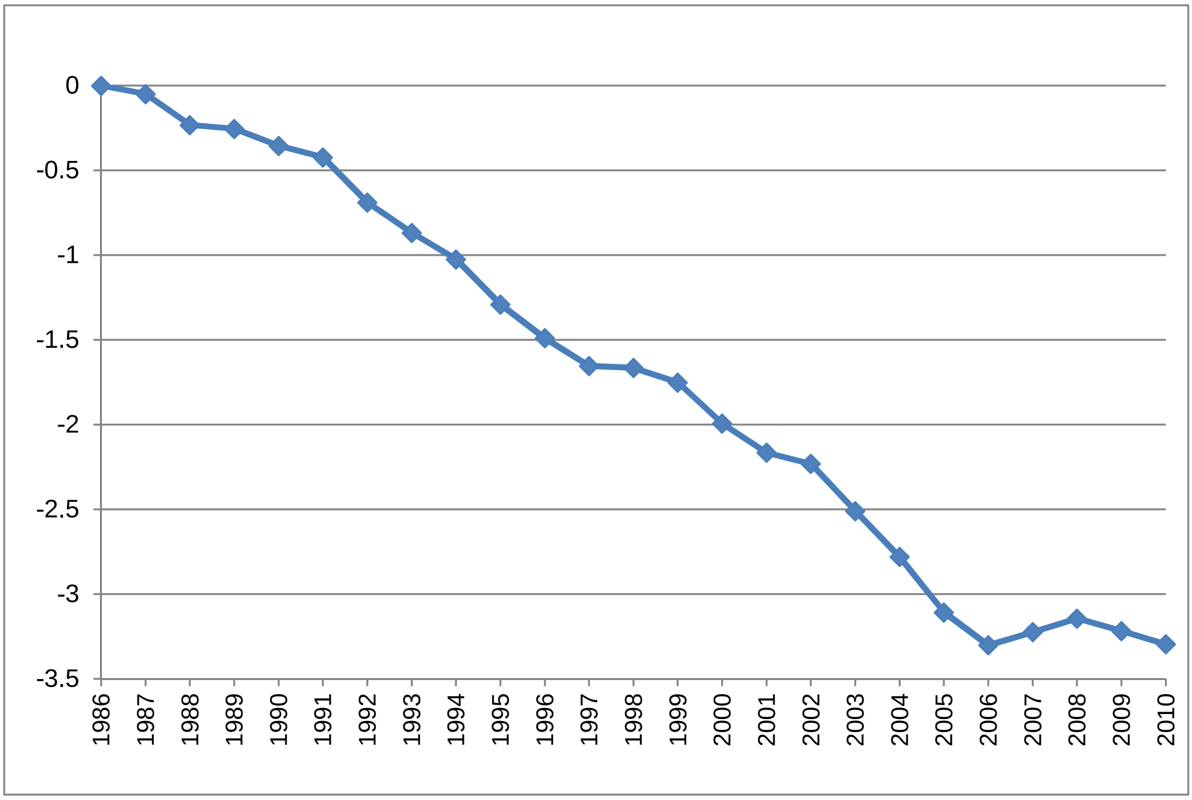


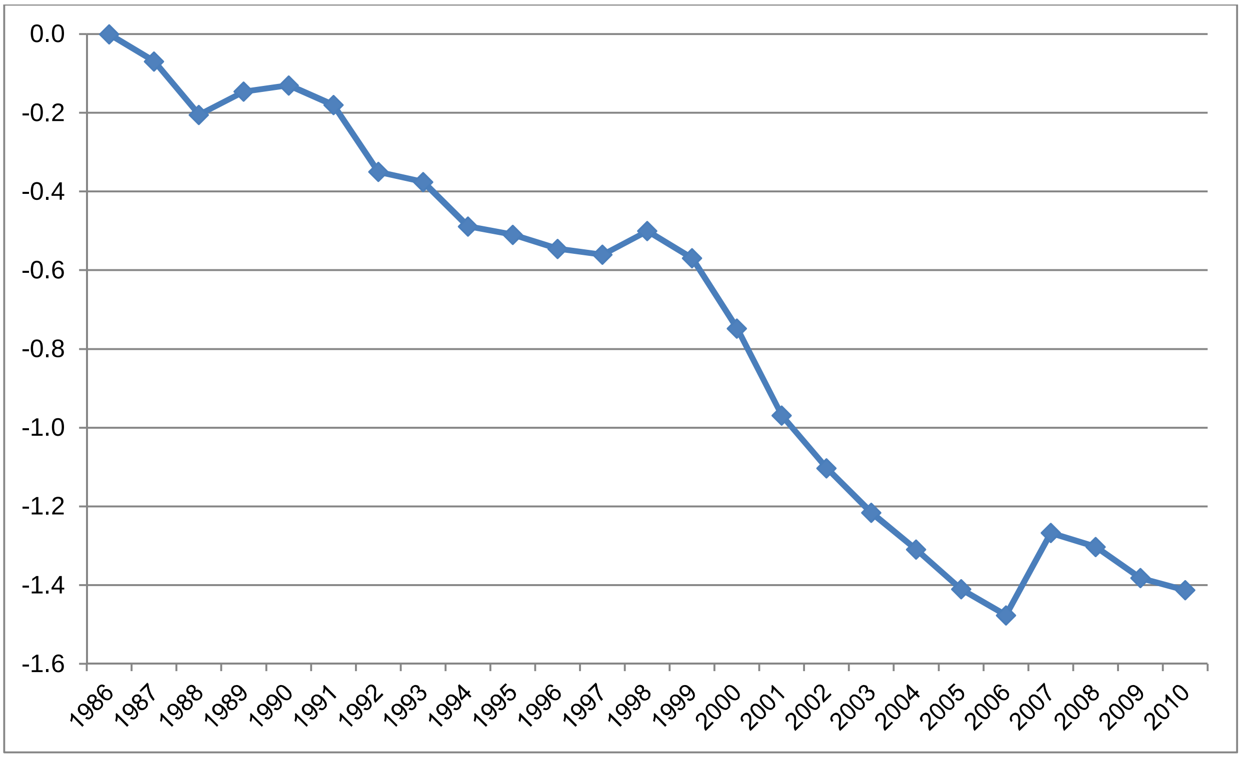
**Figure 3 – Intercept estimates for each year – accounting and market data model**

This figure shows the intercept estimates form the ordered Probit model using accounting and market data for the sample years 1984-2010.



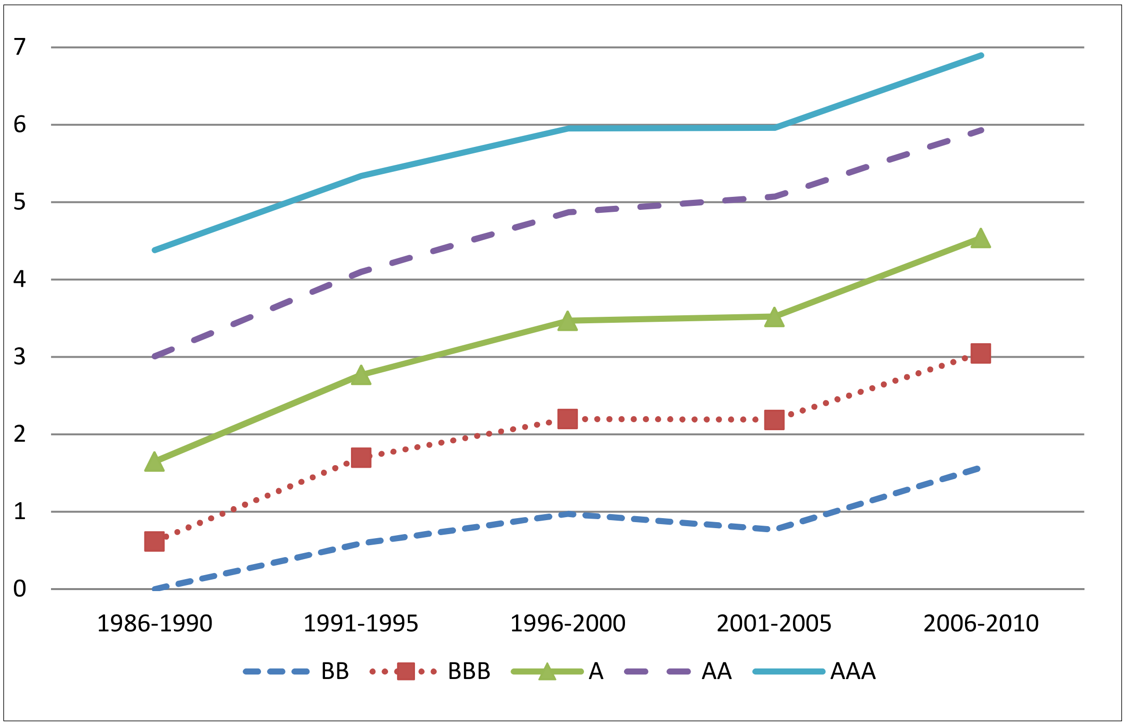
## Figure 4 – Intercept estimates for each year – ‘Distance to Default’ model

This figure shows the intercept estimates form the ordered Probit model using ‘Distance to Default’ as explanatory variable for the sample years 1984-2010.



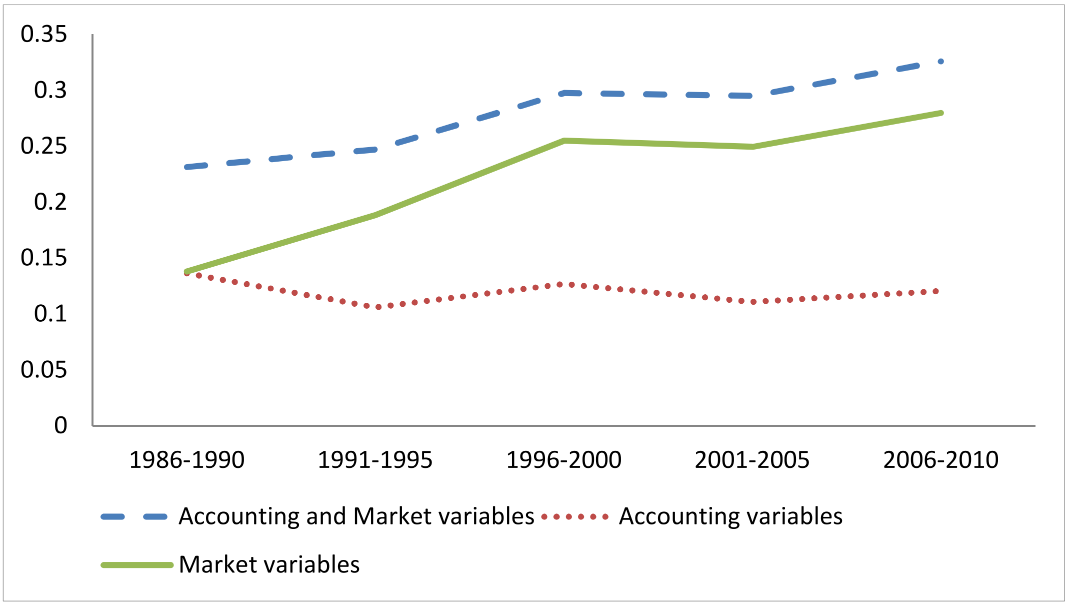

## Figure 5- Estimates of the rating lower boundaries in selected periods

This figure shows the estimates of rating lower boundaries form the ordered Probit model using accounting and market data for several sub-periods.



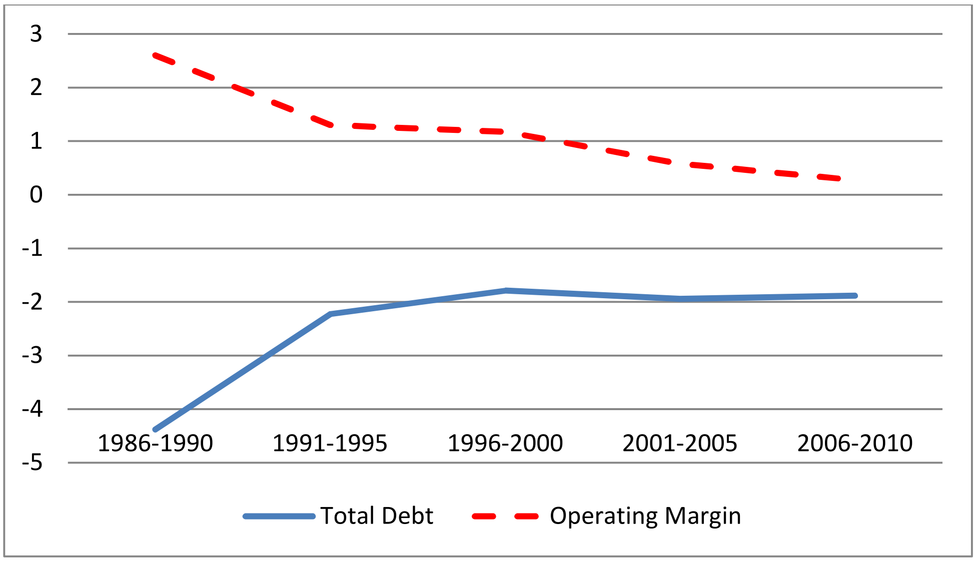
## Figure 6 – Pseudo R-squared in selected periods

This figure shows the Pseudo R-squared of three types of models (Accounting data model, Market data model and Accounting and market data model) in five selected sub-periods.



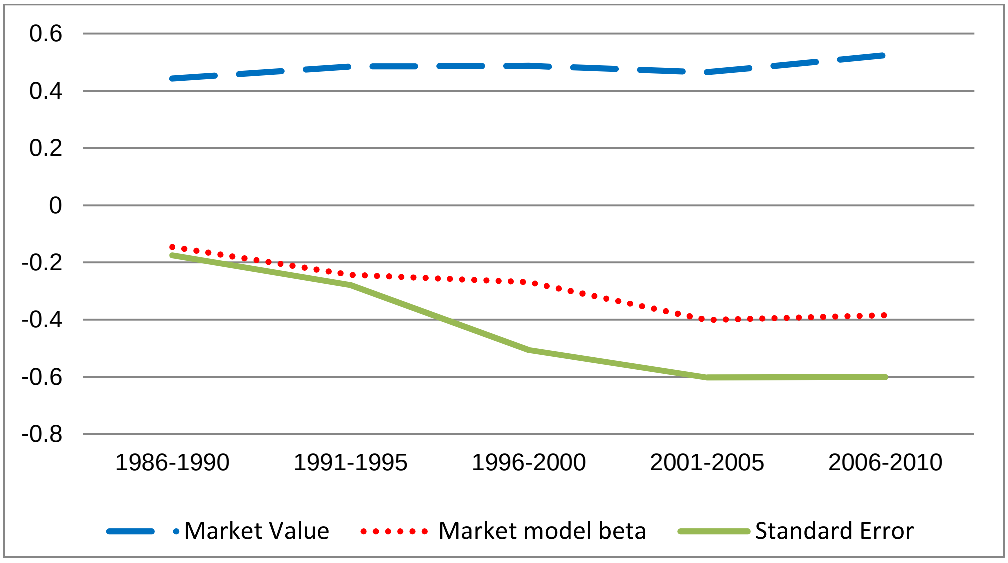

## Figure 7 – Coefficient estimates for 'Total debt leverage' and 'Operating margin'

This figure shows the coefficient estimates for 'Total debt leverage' and 'Operating margin' form the ordered Probit model using accounting and market data in five selected sub-periods.



## Figure 8 - Coefficient trend for ‘Market model beta’, ‘Market model standard error’ and ‘Market value’

This figure shows the coefficient estimates for ‘Market model beta’, ‘Market model standard error’ and ‘Market value’ form the ordered Probit model using accounting and market data in five selected sub-periods.




## Figure 9 – AUC over time

This figure shows the Area Under Curve (AUC) of the Receive Operating Characteristic (ROC) curve for prediction of default using ratings as predictors in the years 1985-2013 (see Table 8). The dashed curves mark the 90% confidence interval [5%, 95%].