# Examining Split Bond Ratings: Effect of Rating Scale

Krishnan Dandapani
Florida International University

Edward R. Lawrence
Florida International University

*In this study we confirm earlier theoretical research findings that differences in the rating scale are responsible for split ratings. We also document that rating scale differences are not the sole explanation for split ratings. The methodology, in measuring the magnitude and documenting of the problem, is new. We first develop an illustrative example by creating two different hypothetical grading scales across two universities and find the patterns in student grades at the two universities. We then compare these results with bond ratings to find if similar patterns exist. The prior empirical research on bond ratings provides strong validation to our analysis. We show that about one-third of the bond split ratings are due to the differences in ratings scales, while the remaining two-thirds are due to the other reasons.*

## Introduction

Investors use bond ratings to measure the riskiness of bonds, and they accordingly make their investments in a firm's securities. Firms are influenced by bond ratings, as the ratings affect the firm's access to capital and its cost of capital. Two major agencies, Moody's Investors Service (Moody's) and Standard & Poor's Ratings Group (S&P), dominate the market in rating publicly traded bonds. The two rating agencies disagree substantially on the ratings for a particular issue or company.[1] Split ratings, especially at the mid range level, have major financial implications. Regulators restrict many investment firms from investing in securities that do not receive investment ratings from at least two major rating agencies. Simi-

[1] Jewell and Livingston (1998) report that about 17 percent of U.S. industrial bond issues from 1983 to 1993 received different letter ratings from Moody's and S&P. In their study on 4,399 bond issues for the time period 1983 to 1993, Cantor, Packer, and Cole (1997) find 54.7 percent of split ratings, mainly consisting of single notch difference (42.1 percent).


0747-5535/07/1300/0065/$2.50

larly, under Basle II agreements, the capital adequacy requirements for bonds may vary depending on the ratings. While examining the financial implications of bond ratings, Billingsley et al. (1985) and Liu and Moore (1987) argue that investors pay more attention to the lower of the two ratings. Hsueh and Kidwell (1998) and Reiter and Ziebart (1991) contend that the higher of the two ratings sets market prices, whereas Jewell and Livingston (1998) affirm that both higher and lower bond ratings have an impact on bond prices and underwriter fees.

In the research on finding the causes for bond splits, Ederington (1986) argues that split ratings are caused by the random errors of the two rating agencies. Cantor (1994) attributes the differences in ratings to the alternative rating methodologies and the result of the judgmental element in the rating process. Morgan (2002) states that the split rating is due to the asset opaqueness of some of the firms, whereas Jewell and Livingston (1999) and Livingston, Naranjo, and Zhou (2006, 2007) attribute the split rating to informational asymmetry and state that the disagreements between the rating agencies are a symptom of firm opaqueness. Boot et al. (2006) provide a novel rationale for credit ratings. They show that credit ratings provide a focal point for firms and their investors and explore the vital contractual relationship between a credit rating agency and a firm through its credit watch procedures. Haggard et al. (2006) explore a complementary explanation by examining whether firm information quality (the precision of signals from the firm's financial information system) is associated with credit rating splits. Using ratings for a sample of new debt issues, they find a positive association between measure of information quality and credit rating analyst disagreement. Their findings are robust in consideration of issuing firm asset opacity.

Although Ederington (1986) does not find any evidence that split ratings result from differences in rating standards (different cutoff points) or weights attached to rating determinants, subsequent studies find that rating determinants or weights attached to rating determinants differ across rating agencies. Moon and Stotsky (1993) examine municipal bond ratings and find that split ratings appear to reflect differences in the weight attached to specific determinants of the ratings and differences in the cutoff points. Pottier and Sommer (1999) study the insurer rating and find evidence that rating determinants and their weights differ across rating agencies.

Horrigan (1966) and West (1970) were the first to predict bond ratings based on the characteristics of the bonds and the issuing firms. These studies assume that the dependent variable is categorized into equally spaced discrete intervals, i.e. the risk differential between an Aaa and Aa rated bond is the same as between a Ba and a B rated bond. Kaplan and Urwitz (1979) argue that while bond ratings convey ordinal information (an Aaa bond is more secure than an Aa bond which is more secure than an A rated bond), these ratings can not be interpreted as equal intervals on a scale. Kaplan and Urwitz further aver that "it is unclear what effect this misspecification has on the Horrigan's and West's studies." Cantor (1994) argues that many of the

differences in the ratings may be due to systematic differences among agencies in determining the acceptable level of risk in any rating category. Subsequently, Cantor and Packer (1997) examine corporate debt ratings and conclude that the differences in ratings reflect differences in rating models.[2] Cantor and Packer (1997) posit:

> Our results suggest that observed differences in average ratings reflect differences in rating scales. They also call into question financial industry regulations that assume equivalence of rating scales. If these results are confirmed by other studies, they should prove useful in the ongoing efforts of the Securities and Exchange Commission to improve its methods of incorporating ratings into its regulations. (p. 1416)

This study provides the theoretical underpinnings for the findings of Cantor and Packer (1997) and demonstrates that one of the reasons for the split rating in bonds is the difference in the rating scale used by the rating agencies.[3] To illustrate this point we create two different hypothetical grade scales across two universities. Then we examine the case of differences in grading scales across the two universities. We find the patterns in student grades due to differences in university grading scales and document its overall impact. Then we compare and contrast it with the bond ratings to find if similar patterns exist. While we would have preferred to compare the actual rating stratum and models used by the rating agencies (such as S&P's and Moody's), these models are proprietary and not readily available in the public domain. Therefore, we use student grades as an illustrative substitute because the design and configurations behind student grading is similar to the bond ratings.

Universities in the United States use letter grades to indicate performance in a course. Based on performance in the courses in a program, the student is awarded an overall grade point average (GPA). This GPA is used in many important scholastic decisions such as retention in school, determining eligibility for a particular major, graduation from a program, as well as in determining meritorious performance (scholarship, employment eligibility) and academic achievements and honors (magna cum laude, cum laude). Different universities have different grade intervals and strata of classification that may lead to situations where students with identical potential and the same performance may be awarded different overall ratings by the two universities. This differing interval grade design enables comparison of the academic grading of students at two different universities with the bond ratings by Moody's and S&P.

As identified in prior research, if asset opaqueness or informational asymmetry is the only cause of bond split rating, then over time this informational asymmetry

[2] We summarize the literature for split ratings in Appendix 1.

[3] Because of the general unavailability of detailed studies on bond split ratings, this study focuses on Cantor, Packer, and Cole's (1997) data on bonds detailed in their Table 1. Where relevant, our study provides a comparative discussion of other research.

between the rating agencies should be eliminated as information inefficiencies disappear. In due course the two ratings should converge. But if the bond rating split is due to the difference in intervaling or rating scale, then the split ratings should be permanent (i.e., the ratings should not converge over time). Livingston, Naranjo, and Zhou (2006, 2007) report that about two-thirds of split rated bonds remain split even after four years from the initial offering, thus confirming that scaling difference is one of the reasons for split ratings along with the other reasons cited in the literature.

To illustrate, we postulate that the probability of getting a B or higher grade is greater in one university whereas the probability of getting a B- or lower is greater at another. We also postulate that the difference in the probability mass function is prominent for the grade categories B, B-, C, and C-; there are marginal differences for the grades B+, C+, and D+; whereas almost no difference exists for the grades A, A-, and D. We argue that if there are measurable scale differences in rating agencies, then one rating agency should rate some companies better in some categories and it should give a lower rating to companies in another category. The percentage of split ratings should be low for the categories at the extreme ends of the scale, whereas the percentage of split ratings should be higher for the categories classified in the middle.

Prior empirical studies on bond ratings confirm these claims that indicate difference in rating scales as one of the probable causes for split ratings in bonds. We also argue that if the rating categories are the same across rating agencies, then the split differences should not be more than one notch due to the differences in rating scales. This argument is based on the comparison with the student grades where we do not find more than one grade split (due to the difference in grading scales). The finding indicates that more than one notch split probably is due to the asymmetric information, judgmental differences, or randomness. If so, these (more than one notch) ratings should converge with time, and we should observe large percentage rating changes for more than one notch splits. Our argument is validated by Livingston, Naranjo, and Zhou (2006, 2007) who show that more than one notch split has the highest percentage of rating changes.

Our comparison to college grading indicates that differences in rating scales can be responsible for about 19 percent of single notch splits, while the remaining bond splits may be due to the other reasons: information asymmetry, judgmental differences, and randomness. Hence, our study not only confirms that differences in rating scale are responsible for split ratings, but also demonstrates that scale differences only provide a partial explanation for the split ratings. Thus, our study underscores that other factors are also responsible for split ratings observed in bonds.

## Biases in Grading Regimes

Assuming that the bias of estimators is non-existent (the grading by professors is unbiased) the bias in grading regimes is due to the difference in the grading models

and intervaling scales used. Three major models have been used in the literature. These are briefly:

(1) The 'Criterion Referenced Standard' where in absolute standards for students performance are pre-determined and compared,
(2) 'Norm Referenced Standard' which focuses on comparison among students and emphasizes comparative grading and
(3) 'Self Referenced Standard' which is noncompetitive and customized and indicates learners' gains based on initial abilities.

Once the learning objectives are delineated using the model, a grading rubric and performance standard is developed. We postulate that whichever standard is used in grade determination, unless an exact and identical system is implemented across universities, it would lead to rating differences. For a student, the number of grading intervals plays an important part in deciding how accurately the grading regime assesses performance. As the number of levels (corresponding to the number of letter grade) increases, the error between the grade point assigned and the true performance level decreases. This error, from a student's point of view is either a loss or a gain depending on how close the student is to the nearest GPA level. The same effect is achieved by increasing the number of courses keeping the number of letter grades fixed. With addition of each course and the subsequent averaging, new GPA levels are introduced. We use *Combination Mathematics* to compute the number of levels. Given a set of all possible unique elements, a combination is a subset where the order of the elements is not important. The number of combinations of size k from a set with n elements is represented as:

$$C_k^n = \frac{n!}{k!(n-k)!}$$

In general, the number of levels for *n* courses and *g* letter grades is given by:[4]

$$\text{Number of levels} = C_n^{(g+n-1)} \tag{1}$$

We assume that the courses are $n_1, n_2, ...., n_n$ and the letter grades are $g_1, g_2, ...., g_g$. After completing all the n courses a student obtains $g_1$ grades in $m_1$ number of courses, $g_2$ grades in $m_2$ number of courses, and so on, so that

$$m_1 + m_2 +, ...., + m_g = n \tag{2}$$

[4] We are calculating in how many different ways we can obtain the grades for a student. We are computing this using equation (1) for the number of combinations of n courses and g grades when a student takes n courses. For example for two courses and two grades (A and B), using the equation the number of levels is three which is: A and A, A and B, and B and B. For two courses and three grades (A, B, and C) the number of levels is six which is: A and A, A and B, A and C, B and C, B and B, and C and C.

The probability of a particular combination $m_1, m_2, ,...., m_g$ is then given by[5]

$$p(m_1, m_2, ...., m_g) = \frac{n!}{m_1! \times m_2! \times, ...., \times m_g!} p_1^{m_1} p_2^{m_2} ....p_g^{m_g} . \quad (3)$$

In the above equation $p_1^{m_1}$ is the probability of observing grade $g_1$ in $m_1$ occasions. We assume that scoring any grade is an equally likely event. As there are g grades, the probability of getting any grade $g_k$ is simply

$$p_{g_k} = 1/g$$

The probability of getting $g_k$ grades on $m_1$ occasions is

$$p_{g_k}^{m_1} = (1/g)^{m_1} . \quad (4)$$

Substituting equation (4) in equation (3) we get

$$p(m_1, m_2, ...., m_g) = \frac{n!}{m_1! \times m_2! \times, ...., \times m_g!} \left(\frac{1}{g}\right)^{m_1} \left(\frac{1}{g}\right)^{m_2} .... \left(\frac{1}{g}\right)^{m_g}$$

$$p(m_1, m_2, ...., m_g) = \frac{n!}{m_1! \times m_2! \times, ...., \times m_g!} \left(\frac{1}{g}\right)^{m_1 + m_2 + ..... m_g} . \quad (5)$$

Substituting equation (2), the above equation can be written as

$$p(m_1, m_2, ...., m_g) = \frac{n! g^{-n}}{m_1! \times m_2! \times, ...., \times m_g!} . \quad (6)$$

We assume two grading regimes using the same number of letter grades g such that $w_{13}$ is the numerical weight assigned by regime 1 to the letter grade $g_3$ .We can write the GPA of a student with the $m_1, m_2, ,...., m_g$ combination under regime 1 as:

$$GPA_1 = (m_1 \times w_{11} + m_2 \times w_{12} +, ...., + m_g \times w_{1g}) \quad (7)$$

The GPA thus obtained can be converted into the letter grades, and the student can be given an overall letter grade. Thus, it is possible to compute the probability of obtaining a final letter grade (corresponding to a grade combination) using equation (6) and to plot the probability mass function of the regime. If we now compare two grading regimes each with its own spacing of intervals (its own assignment of weights $w_{ij}$ to the letter grades), then statistically, we are comparing two random variables with their own mass functions.

[5] The formula for the computation of probability in equation (3) can be obtained from any book on probability. See Hadley (1967, p. 309).

In U.S. universities, a student passing a course is awarded one grade of typically ten possible letter grades (A, A-, B+, B, B-, C+, C, C-, D+, and D).[6] Uniformly an A grade carries a weight of 4 points, a B grade has a weight of 3 points, and a C grade has a weight of 2 points, while 1 point weight is given for a D grade. The differences emerge from the treatment of plus and minus grades (where they exist). The plus and minus grades usually have different weights assigned in different universities. One university may assign a weight of 3.75 for an A- grade, while another university may assign it a weight of 3.66. In Panel A of Appendix 2, we provide the grade point scale for some of the universities in U.S.A. (Florida State University, Northwestern University, and University of Minnesota). In Panel B of Appendix 2, we present the grading standards of three universities in the same university system in Florida (University of Florida, Florida State University, and Florida International University), and these are very different.[7] Typically an undergraduate student is required to take 40 courses, and a master's student is required to take 10 or more courses to obtain a degree. The student is awarded a grade in each of the courses. Employing those grades, an overall GPA is calculated, which can then be expressed as a letter grade.

For illustrative purposes we create two hypothetical grade scales in Table 1 for two universities, Alpha and Beta. Note that these grade scales are generated and do not correspond to any of the universities that we show in Appendix 2. We do not use the grading scales as reported in Appendix 2 as they are too symmetrical and do not yield the grade splits that we show through this example. In particular, grade splits are dependent on the grade scales used. For our illustrative purposes, however, the important point here is that different grade scales can result in split grades. As can be readily seen, the grading intervals between the two universities differ.

**Table 1—Grading Scales at Two Universities**

| | A | A- | B+ | B | B- | C+ | C | C- | D+ | D |
|---|---|---|---|---|---|---|---|---|---|---|
| University Alpha | 4 | 3.7 | 3.5 | 3 | 2.7 | 2.5 | 2 | 1.7 | 1.5 | 1 |
| University Beta | 4 | 3.6 | 3.4 | 3 | 2.6 | 2.4 | 2 | 1.6 | 1.4 | 1 |

As an illustration, if a student at university Beta taking ten courses, gets an A in two courses, A- in two courses, B+ in two courses, B- in one course , C in two courses and a C- in one course, using equation (7) we can compute her GPA:

$$\text{GPA}_{\text{Beta}} = (2\times 4/10 + 2\times 3.6/10 + 2\times 3.4/10 + 1\times 2.6/10 + 2\times 2/10 + 1\times 1.6/10)$$

[6] Universities give letter grades D- and F also, but for comparative purposes with the number of rating categories (ten), these grades are omitted from our discussion.

[7] Many universities have only four classifications (A, B, C, and D) and other universities such as the University of Nebraska–Lincoln, the University of Minnesota, and the Florida International University follow the same grading scale. Across these universities the bias in student performance due to grading scale won't exist.

$$GPA_{Beta} = 3.02$$

The weights used in the above equation are the individual grade weight divided by the number of courses. For converting the GPA score to letter grade, we use the following table.

**Table 2—Conversion Table to Get Letter Grades from GPA**

| Alpha | Beta | Grades |
|---|---|---|
| GPA = 4 | GPA = 4 | A |
| 4 > GPA ≥ 3.7 | 4 > GPA ≥ 3.6 | A- |
| 3.7 > GPA ≥ 3.5 | 3.6 > GPA ≥ 3.4 | B+ |
| 3.5 > GPA ≥ 3 | 3.4 > GPA ≥ 3 | B |
| 3 > GPA ≥ 2.7 | 3 > GPA ≥ 2.6 | B- |
| 2.7 > GPA ≥ 2.5 | 2.6 > GPA ≥ 2.4 | C+ |
| 2.5 > GPA ≥ 2 | 2.4 > GPA ≥ 2 | C |
| 2 > GPA ≥ 1.7 | 2 > GPA ≥ 1.6 | C- |
| 1.7 > GPA ≥ 1.5 | 1.6 > GPA ≥ 1.4 | D+ |
| 1.5 > GPA ≥ 1 | 1.4 > GPA ≥ 1 | D |

In the above example the GPA of 3.02 is equivalent to an overall letter grade of B. With the assumed scale we compute[8] the probability mass function for four and ten courses using equation (6). We also assume that all possible combinations of grades are equally likely. As we are making comparison with the bond ratings where ratings are given for all characteristics simultaneously, we assume that student takes all courses at the same time or they get the grades for different courses at the same time, i.e., grades depend only on the performance, not capabilities.

In Figure 1A we plot the probability mass function for four courses, and in Figure 1B we plot the probability mass function for ten courses. From the figures it is evident that the probability for obtaining any overall GPA letter grade is different across the two universities. Though the percentage difference falls, the distinction between the universities does not disappear as the number of courses increases. The loss or gain arising from the differences in grading scales is a relative loss or gain for the students and can have severe implications for the student.

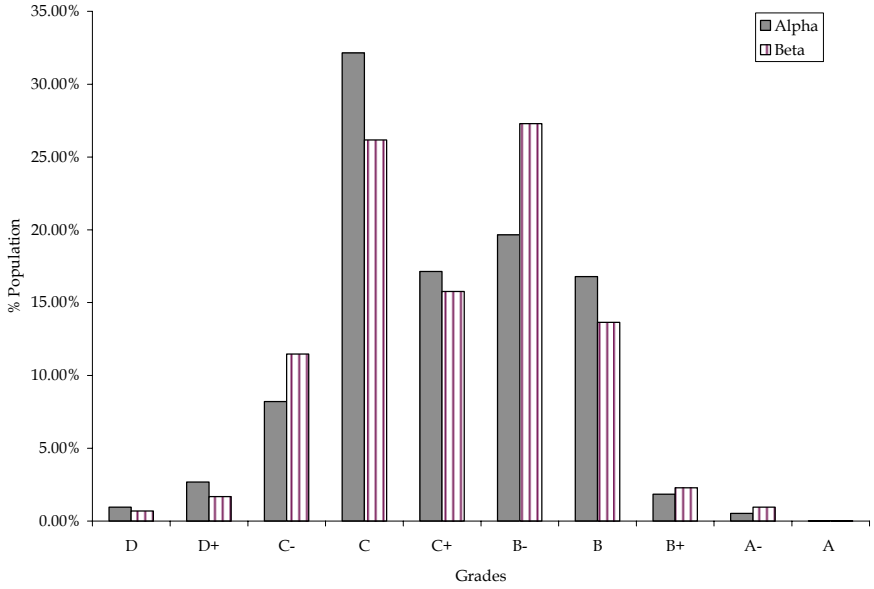
From the mass functions for ten courses, we plot the difference in the mass functions of the two universities for each grade level in Figure 2. The difference is prominent in the grade categories B, B-, C, and C-. There are marginal differences for the grades B+, C+, and D+. The differences in grades A- and D are almost trivial. The probability of getting a B and C grade is higher in university Alpha, whereas the probability of getting grades B- and C- is higher in university Beta. The probability

[8] The computational intensity increases dramatically with an increase in the number of courses. With ten grades (as shown in Table 1) the total number of combinations for four courses is $10^4$ which increases to $10^{10}$ for ten courses and $10^{40}$ for 40 courses.

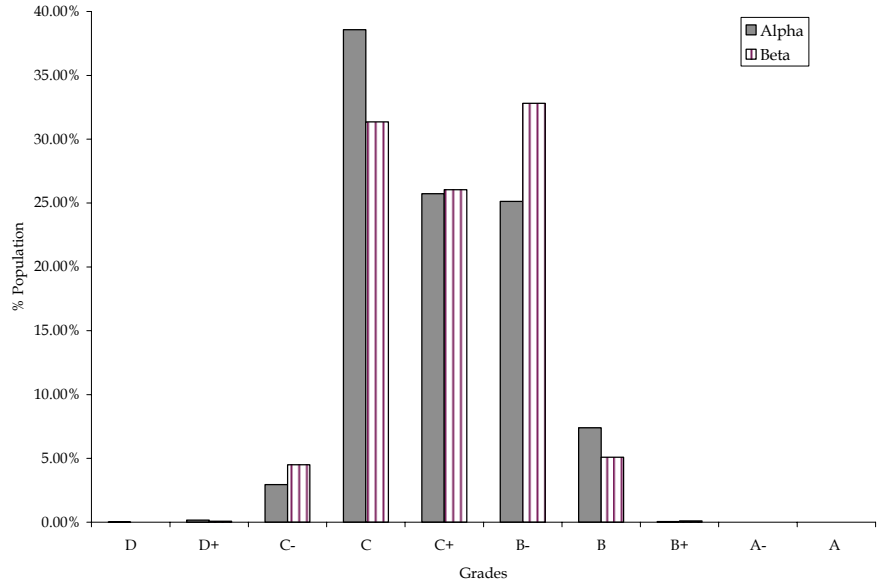
**Figure 1—Probability Mass Function**

Figure 1A shows the probability mass function for four courses, and Figure 1B shows the probability mass function for ten courses. The figures show that the probability for obtaining any overall GPA letter grade is different between the universities. Although the percentage difference declines, the distinction between the universities does not disappear as the number of courses increase.

Panel 1A (for four courses)




Panel 1B (for ten courses)

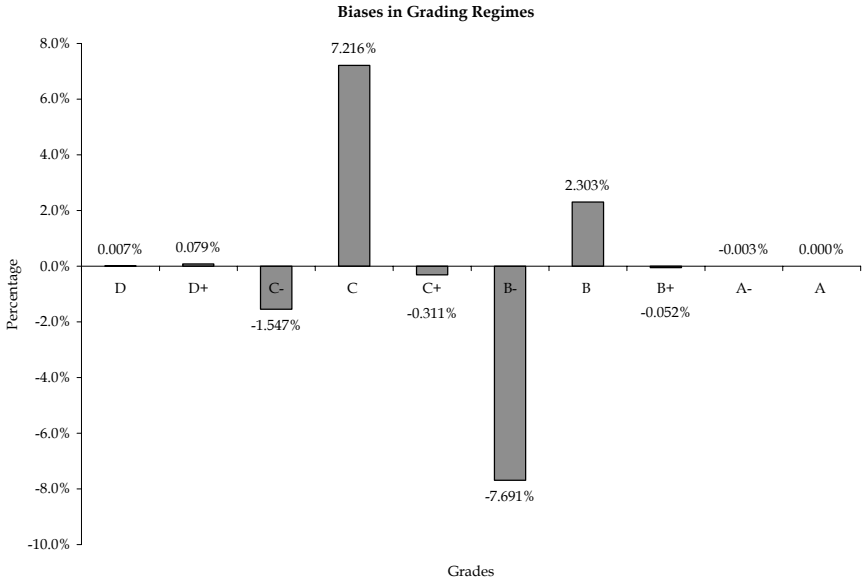


of getting B and higher grades is higher in university Beta, whereas the probability of getting B- and below grades is higher in university Alpha.

We observe only single grade (notch) differences across the two universities. With four courses, the population with split grades is 23.54 percent, for five courses the population with split grades is 20.70 percent, and for ten courses the population with split grades reduces to 19.20 percent.

**Figure 2—Difference in the Mass Function of Two Universities**

In the following figure we plot the difference in the mass functions of the two universities for each grade level for ten courses. We observe prominent differences in the grade categories B, B-, C, and C-, marginal differences for the grades B+, C+, and D+, whereas the differences in grades A- and D are almost trivial. Probability of getting B and C grades is higher in university alpha whereas the probability of getting grades B- and C- is higher in university beta. The probability of getting B and higher grades is higher in university beta, whereas the probability of getting B- and below is higher in university alpha




## The Rating Methodology

As stated by S&P, "A credit rating is S&P's opinion of the general creditworthiness of an obligor, or the creditworthiness of an obligor with respect to particular debt security or other financial obligation, based on relevant risk factors."[9] For Moody "ratings are opinions of future relative creditworthiness, derived by fundamental credit analysis and expressed through the familiar Aaa-C symbol system. Fundamental credit analysis incorporates an evaluation of franchise value, financial statement analysis, and management quality. It seeks to predict the credit perform-

[9] S&P Corporate Ratings Criteria, 2006, p. 8.

ance of bonds, other financial instruments, or firms across a range of plausible economic scenarios, some of which will include credit stress."[10]

The rating process generally includes quantitative financial analysis of the company based on financial reports, qualitative analysis of the management, and legal analysis including regulatory changes and labor relations. The rating process generally includes three types of information; 1) publicly available information which includes audited financial statements and qualitative information such as media reports on the state of the firm and the industry, 2) information disclosed by the issuer themselves which includes financial information on the operating position of the firm, information on accounting policy, management skills, competitive positioning, and the corporate strategy of the firm, 3) the information provided by the disgruntled former employees or by competitors. The final ratings usually are based on the judgment of the rating evaluators. In this way, the bond ratings are analogous to the student grades as a professor typically incorporates subjective judgment in determining the final grade to be awarded. Normally this subjective judgment is regarding the student's participation and interest in the class over and beyond what the grading rubric suggests.

In Table 2 we present the interpretation of various credit ratings issued by S&P and Moody's. There are ten ratings classes for S&P, from AAA down to D and nine rating classes for Moody, from Aaa to C. Thus, the differential strata mechanism translates to differential scales and intervals in bond ratings. The ratings above BBB for S&P and Baa for Moody are classified as Investment Grade, and the ratings below these are classified as Speculative Grade.

**Table 3—Debt Rating Symbols and their Interpretations**

| Explanation | Standard & Poor's | Moody's Service |
|---|---|---|
| Highest Grade | AAA | Aaa |
| High Grade | AA | Aa |
| Upper Medium Grade | A | A |
| Medium Grade | BBB | Baa |
| Lower Medium Grade | BB | Ba |
| Speculative | B | B |
| Poor Standing | CCC | Caa |
| Highly Speculative | CC | Ca |
| Lowest Quality, No Interest | C | C |
| In Default | D | |

## A Comparison of Student Grades and Bond Rating

The bond ratings by the agencies are similar to the cumulative grade point average score given to a student at a university.[11] A bond issued by a corporation is rated

[10] *Understanding Moody's Corporate Bond Ratings and Rating Process*, 2002, p. 5.

based on various different criteria. Depending on the company's performance in each criterion, a final rating is determined. This is analogous to a situation where two students in the same field may get similar grades in all their courses at two different universities, but may end with different cumulative GPAs; one may have an overall letter grade of A- while the other may receive a B+. Similar to student grades, S&P and Moody's have several letter ratings. Similar to the grading of students at two universities, the differing intervals and varied rating scales can lead to different ratings by S&P and Moody's. As our purpose is to examine the effect of the differences in scales on the rankings of bonds, we assume all grades are equally likely for a student. This way, we are separating the differences in capabilities that are dependent on students from the differences in grading scales which are dependent on the measuring system.

Our grading scale results suggest that the difference in scale is one of the main causes behind the split ratings. Based on our results of the different grading scales in the two universities, answers to the following three significant questions indicate that the difference in scale is one of the causes behind split ratings of bonds.

*Question 1:* Do ratings converge over time thereby eliminating the split ratings?

If the difference in scale is the cause behind split ratings, then the split ratings should be permanent (i.e., the ratings should not converge with time). If the split ratings are due to random error (Ederington, 1986) or due to the assets opaqueness[12] of some of the firms (Morgan, 2002) or due to informational asymmetry (Livingston, Naranjo, and Zhou; 2006, 2007) only, then in a few years, the split-rated bonds should converge to the same ratings as the information inefficiencies disappear. Livingston, Naranjo, and Zhou (2006, 2007) report that about two-thirds of split-rated bonds remain split four years after an initial offering. This confirms that scaling or interval difference is one of the reasons for split ratings along with the other reasons cited in the literature.

*Question 2:* Do some ratings have greater chances of divergence than other ratings?

The first four categories of bond ratings are considered investment grade, and the remaining categories are considered speculative grade. Comparing with student

[11] In real life the underlying student populations differ in grade distributions, whereas bond ratings are available for the same underlying firm populations. As we are comparing student grading to bond ratings, to make the comparison more robust, we have assumed similar populations for the grade distributions.

[12] As opaqueness may be inherent in the assets, it may not be eliminated in years as information embedded in assets fails to clearly materialize. This opaqueness should not be different for different rating agencies. Here our argument is about the informational asymmetry between rating agencies where one rating agency has more information than another, leading to the difference in the rating. If the split is only due to the informational asymmetry, then over time this information asymmetry will disappear. Hence, the ratings should converge.

grades, the grades A, A-, B+, and B are equal to the investment grade category. The remaining are in the speculative category. In Figure 2 we observe that the probability of getting B and higher grades is higher in university Beta, whereas the probability of getting B- and below is higher in university Alpha. If there is a scale difference in rating agencies, then one rating agency should rate some companies better in some categories whereas it should give a lower rating to companies in some other category. Consistent with our findings, Packer and Cole (1997) state that when agencies disagree over rating assignments, Moody rates higher than S&P's in only 42.4 percentage of the split-rated investment-grade issues, whereas it rates higher more than 60 percentage of the split-rated non-investment-grade bonds.

*Question 3:* Are the percentage of split ratings low for the categories at the extreme ends of the rating scale whereas the percentage of split rating is higher for the categories lying in the middle?

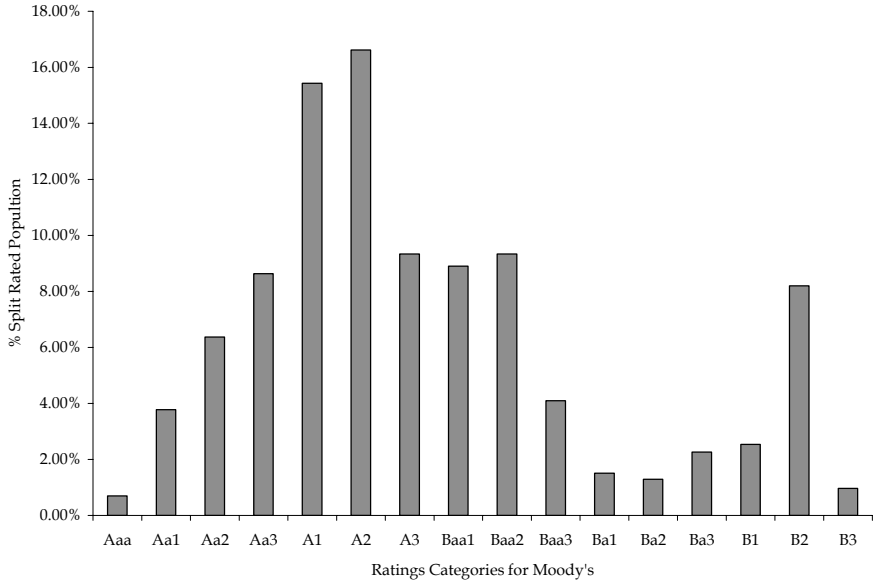
In Figure 2 we observe that the difference in the probability mass function is prominent for the grade categories B, B-, C, and C-; there are marginal differences for the grades B+, C+, and D+; whereas the differences in grades A- and D are almost nonexistent. If the difference in scale is one of the reasons for the split ratings, then the same trend should be observed in the bond split ratings. Cantor, Packer, and Cole (1997) do a comprehensive study of the split ratings of bonds for Moody's and S&P for a total of 4,399 bonds over a ten year period. Their Table 1 has the distribution of Moody's and Standard & Poor's' ratings. We use the data from Table 1 of Cantor, Packer, and Cole (1997). For each rating class we plot the percentage of one notch splits for Moody's and Standard & Poor's. Figure 3 shows the resulting graph which generally matches our findings. We observe high numbers of split ratings in the middle categories of Moody's scale whereas the extreme categories have few split ratings. The only exception is the B2 category which has significantly higher split ratings and may be due to the reasons other than the differences in rating scales.

This discussion indicates that a difference in rating scales is likely one of the causes for split ratings of bonds. In the literature, difference in weights, informational asymmetry, judgmental error, and randomness are other accepted possible reasons for split ratings. None of the previous studies has attempted to quantify and apportion the magnitude of the split due to each of the above mentioned causes. Our study provides bounds for the magnitude of splits due to the differences in rating scales and due to the other reasons for split ratings of bonds.

If the number of rating categories is the same across rating agencies, then the split differences should not be more than one notch due to the differences in rating scales. For our example we observe that under the assumption of the same grading categories, the split in grades is for only one grade notch. Comparing it to the split ratings in bonds suggests that differences in rating scales are responsible only for single notch splits.

**Figure 3—Ratings Categories for Moody's**

Using the data from Table 1 of Cantor, Packer and Cole (1997) we plot the number of one notch split for Moody's. We observe high numbers of split ratings in the middle categories of Moody's scale whereas the extreme categories have few split ratings.




This argument suggests that differences in ratings scales are not responsible for the more than one notch of bond splits. Hence, more than a one notch split must be due to the asymmetric information or judgmental differences and/or due to randomness. These split ratings should converge with time, and we should observe a large percentage rating migration for more than one notch splits. Livingston et al. (2005) confirm this and shows that "with no exception, the sample of bonds with more than one notch split has the highest percentage of rating changes, while the non-split sample have the lowest percentage of rating changes."

For ten grades and four courses we find that there are 23.54 percent split grades; for five courses there are 20.4 percent split grades; and for ten courses there are 19.20 percent split grades. Though the bond rating methodology is proprietary information, there are indications that there are definitely more than four areas evaluated for bond ratings. (We compare them to the courses.) Table 2 shows that there are nine matching categories across S&P and Moody. A comparison to college grading indicates that differences in rating scales can be responsible for a maximum of 23 percent single notch splits. For a more reasonable ten evaluation areas for bond ratings, the differences in rating scales can be responsible for almost 19 percent of single notch splits. Note that these numbers are indicative and are entirely dependent

on the grading scales. Hence, depending on the differences in the rating scales, the percentage split can be higher or lower than 19 percent.

In their study on 4,399 bond issues for the time period 1983 to 1993, Cantor, Packer, and Cole (1997) find 54.7 percent split ratings with a vast majority (42.1 percent) single notch differences. Our study suggests that of the total 54.7 percent of split ratings about one-third could be due to the differences in rating scales and the remaining two-thirds of split ratings are due to the other reasons.

## Conclusion

Cantor and Packer (1997) examine corporate debt ratings and find that the differences in ratings reflect differences in rating models. They conclude that "[i]f these results are confirmed by other studies, they should prove useful in the ongoing efforts of the Securities and Exchange Commission to improve its methods of incorporating ratings into its regulations." In this study we confirm the findings of Cantor and Packer (1997) by developing an illustrative example, creating two hypothetical grading scales across two universities, and finding the patterns in overall GPA grades at the two universities. We then compare these results with bond ratings and show that similar patterns exist. This suggests that one of the reasons for the split rating in bonds is the difference in the rating scale used by the rating agencies. Prior research and empirical results on bond ratings support our findings, indicating difference in rating scales as one of the causes for split ratings in bonds. Our results suggest that about one-third of the bond split ratings can be due to the differences in ratings scales, while the remaining two-thirds are due to other reasons such as information asymmetry, judgmental differences, and randomness.

## References

1. Billingsley, R.S., R.E. Lamy, M.W. Mar, and G.R. Thompson, "Split Ratings and Bond Reoffering Yields," *Financial Management*, 14 (1985), pp. 59-65.

2. Boot, Arnoud, W.A. Todd, T. Milbourn, and Anjolein Schmeits, "Credit Ratings as Coordination Mechanisms," *Review of Financial Studies*, 19 (2006), pp. 81-118.

3. Cantor Richard, "The Credit Rating Industry," *Federal Reserve Bank of New York, Quarterly Review* (Summer-Fall 1994).

4. Cantor, Richard, Frank Packer, and Kevin Cole, "Split Ratings and the Pricing of Credit Risk," Research Paper 9711, Federal Reserve Bank of New York (1997).

5. Cantor, Richard, and Frank Packer, "Differences of Opinion and Selection Bias in the Credit Rating Industry," Journal of Banking & Finance 21 (1997), pp. 1395-1417.

6. Ederington, L., "Why Split Ratings Occur?" *Financial Management*, 15 (Spring 1986), pp. 37-47

7. Hadley G., *Introduction to Probability and Statistical Decision Theory* (San Francisco, California: Holden-Day Inc., 1967).

8. Haggard, K. Stephen, Ravi Jain, Xiumin Martin and Raynolde Pereira, "Information Quality and Split Bond Ratings," working paper (2006).

9. Horrigan J., "The Determination of Long-term Credit Standing with Financial Ratios," *Journal of Accounting Research*, 4 (supplement, 1966), pp. 44-62.

10. Hsueh, P., and D. Kidwell, "Bond Ratings: Are Two Better Than One?" *Financial Management*, 17 (1988), pp. 46-53.

11. Jewell, Jeff, and Miles Livingston, "Split Ratings, Bond Yields, and Underwriter Spreads for Industrial Bonds," *Journal of Financial Research*, 21 (1998), pp. 185-204.

12. Jewell, Jeff, and Miles Livingston, "A Comparison of Bond Ratings from Moody's S&P and Fitch IBCA," *Financial Markets, Institutions and Instruments*, 8, no. 4 (1999).

13. Kaplan, R., and G. Urwitz, "Statistical Models of Bond Ratings: A Methodological Inquiry," *Journal of Business*, 52 (1979), pp. 231-261.

14. Liu, P., and W. Moore, "The Impact of Split Bond Ratings on Risk Premia," *Financial Review*, 22 (1987), pp. 71-85.

15. Livingston, Miles, Andy Naranjo, and Lei Zhou, "Split Bond Ratings and Ratings Migration," working paper, University of Florida (2006).

16. Livingston, Miles, Andy Naranjo, and Lei Zhou, "Asset Opaqueness, Split Ratings and Rating Migration," forthcoming in *Financial Management* (2007).

17. Morgan, D.P., "Rating Banks: Risk and Uncertainty in an Opaque Industry," *American Economic Review*, 92 (2002), pp. 874-888.

18. Moon, C.G., and J.G. Stotsky, "Testing the Differences between the Determinants of Moody's and Standard & Poor's Ratings: An Application of Smooth Simulated Maximum Likelihood Estimation," *Journal of Applied Econometrics*, 8 (1993), pp. 51-69.

19. Pottier, S.W., and D.W. Sommer, "Property-liability Insurer Financial Strength Ratings: Differences across Rating Agencies," *The Journal of Risk and Insurance*, 6 (1999), PP. 621-642.

20. Reiter, S., and D. Ziebart, "Bond Yields, Ratings, and Financial Information: Evidence from Public Utility Issues," *Financial Review*, 26 (1991), pp. 45-73

21. West R., "An Alternative Approach to Predicting Corporate Bond Ratings," *Journal of Accounting Research*, 7 (1970), pp. 118-127.

22. S & P Corporate Ratings Criterion www.corporatecriteria.standardandpoors.com

23. Understanding Moody's Corporate Bond Ratings and Rating Process
http://www.moodys.com

## Appendix 1

| Researchers | Reasons for split ratings |
|---|---|
| Ederington (1986) | Random errors; different rating standards; differing factors and weights. |
| Moon and Stotsky (1993) | Measurement errors/intervaling; differences in the weight attached to specific determinants of ratings and differences in cut off points. |
| Cantor (1994) | Alternative rating methodologies; judgmental element in rating process |
| Jewell and Livingston (1999) | Information asymmetry |
| Pottier and Sommer (1999) | Measurement errors/intervaling; rating determinants and their weights differences. |
| Morgan (2002) | Asset opaqueness |
| Livingston, et al. (2005) | Informational asymmetry; firm opaqueness |
| Boot, et al. (2006) | Control and coordination for credit watch |
| Haggard, et al. (2006) | Information quality and precision of financial signals |

## Appendix 2

In Panel A, we provide the grade point scale for some of the universities in USA. In Panel B, we provide the grade point scale for some of the universities in one state (Florida) of USA. We observe difference in the grading scales both at the country level as well as state level.

Panel A

| | Grade Points | | |
|---|---|---|---|
| Grades | Florida State University | Northwestern University | University of Minnesota |
| A | 4 | 4 | 4 |
| A- | 3.75 | 3.7 | 3.67 |
| B+ | 3.25 | 3.3 | 3.33 |
| B | 3 | 3 | 3 |
| B- | 2.75 | 2.7 | 2.67 |
| C+ | 2.25 | 2.3 | 2.33 |
| C | 2 | 2 | 2 |
| C- | 1.75 | 1.7 | 1.67 |
| D+ | 1.25 | -- | 1.33 |
| D | 1 | 1 | 1 |

Panel B

| | Grade Points | | |
|---|---|---|---|
| Grades | University of Florida | Florida State University | Florida International University |
| A | 4 | 4 | 4 |
| A- | -- | 3.75 | 3.67 |
| B+ | 3.5 | 3.25 | 3.33 |
| B | 3 | 3 | 3 |
| B- | -- | 2.75 | 2.67 |
| C+ | 2.5 | 2.25 | 2.33 |
| C | 2 | 2 | 2 |
| C- | -- | 1.75 | 1.67 |
| D+ | 1.5 | 1.25 | 1.33 |
| D | 1 | 1 | 1 |