Market-Based Evaluation for Models to Predict Bond Ratings and Corporate Bond Trading Strategy

Konan Chan
Department of Finance
College of Management
National Taiwan University
Taipei, Taiwan 10660

Phone: 886-2-23698955
Fax : 886-2-23698955
kchan@mba.ntu.edu.tw

and

Narasimhan Jegadeesh
Department of Finance
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign
Champaign, IL 61820
Phone (217) 333-1505
Fax (217) 244-3102
jegadees@uiuc.edu

November 15, 2001

Market-Based Evaluation for Models to Predict Bond Ratings and Corporate Bond Trading Strategy

Abstract

Previous studies find that most of the bond rating process can be captured by statistical models based on publicly available accounting information. However, these papers use agency ratings as the benchmark to assess the models and ignore the evidence that agency ratings may not be accurate in a timely manner. In this paper, we propose a new approach which incorporates bond returns to evaluate bond rating models. Relative rating strength portfolios, formed by buying under-rated bonds with agency ratings lower than model ratings and selling over-rated bonds with agency ratings higher than model ratings, are employed to test the performance of different statistical models in rating predictions. Our results show that one version of multiple discriminant analysis model can generate statistically significant abnormal returns, suggesting that the ordered probit model which is believed to possess theoretical advantages in classifying bonds does not perform better. The result also indicates that using traditional measures to evaluate models may be misleading. However, the abnormal return from the discriminant analysis model is not economically significant, with only 5% over a 5-year horizon. This implies that the corporate bond market is relatively efficient in processing the accounting information

.

I. Introduction

The most important concern for corporate bond investors is issuers' creditworthiness, or the ability of issuers to pay the scheduled interests and principal on time. The reason that the creditworthiness of issuers is critical for investors is due to its significant impact on the pricing of bonds. The general perception is: the less creditworthiness of an issuer, the higher the default possibility in the future, and the higher the expected return investors would demand, which in turn results in a lower current price for the issue. Therefore, if investors can identify the default risk of an issue beforehand, they can price the issue with better accuracy as well as to ensure the quality of their investment portfolio.

Since the default probability is unobservable and may depend on the firm's business, industry, financial status, as well as many other factors, it's not an easy task for investors to evaluate the default risks of thousands of bond issues in the market. As a result, the professional rating services supplied by major rating agencies, such as Moody's Investor Service (Moody's) and Standard and Poor's (S&P), become important evaluation criteria for investors to assess default risks. In fact, some studies indicate that bond ratings proxy for the default risk of firms, and are highly correlated with the bond yields to maturity. Hickman (1958) pioneers this research work and finds a monotonic relationship between ratings and default rates during the period of 1900-1943. Later, a series of bond mortality research by Altman (1989, 1990, 1992) generally supports Hickman's idea that the lower the rating, the higher the default risk in terms of probability and dollar amount. Moreover, Bennett, Esser and Roth (1994) shows that the lower the rating, the higher the bond yield spread over the Treasury yield during the period of 1983-1992. All these results imply that ratings can be used as surrogates of firms' long-run credit risks and investment quality, a statement claimed to be true by Moody's and S&P.

Due to the importance of default risks and the close relationship between default risks and bond ratings, a bulk of literature have focused on the rating prediction based on a limited number of accounting variables. Horrigan (1966) starts this line of research by using a simple ordinary least squares (OLS) regression model. He finds that more than half of agency ratings can be correctly replicated with the use of six variables in his model. Pinches and Mingo (1973, 1975) apply multiple discriminant analysis (MDA) model to improve the statistical fit. Ederington (1985) further examines the performance of different models in bond rating predictions and concludes that ordered probit and multinomial logit models dominate OLS and MDA models. The general finding of these studies is that most of the statistical models examined correctly replicate about two-thirds agency ratings of bonds. This suggests that the proposed models can capture the essence of the rating process of rating agencies.

However, all these papers use agency ratings as the benchmark to evaluate the performance of statistical models and ignore the evidence that agency ratings may not be accurate in a timely fashion. Specifically, Weistein (1977) finds that bond price changes are fully anticipated during the period from 6 to 18 months before the bond rating changes, a result further supported by Pinches and Singleton (1978) which examine the corresponding stock market returns. More recently, Holthausen and Leftwich (1986) and Hite and Warga (1997) detect significant announcement effects of bond downgrades on stock and bond markets, respectively. The abnormal return within 6 to 12 months prior to rating change announcements is very significant, about two to ten times larger than the abnormal return in the announcement month, for both downgrades and upgrades. Wansley and Clauretie (1985) also document a similar results for bonds with downgrades. In other words, the market has anticipated most of the rating process long before the rating change announcements, suggesting that ratings provided by major rating agencies may not be unbiased estimators of credit risks of firms at each point of time. Accordingly, previous rating prediction studies falsely assume that agency ratings are true ratings and do not fairly evaluate the statistical models they develop.

One possible reason why agency ratings may be biased is that Moody's and S&P cannot afford a day-to-day close monitoring for tens of thousands of corporate bond issues in the market. For example, Ederington and Yawitz (1987) survey the bond rating process and find that there are only 150 security analysts for S&P and 80 for Moody's to rate corporate bond issues. Although both rating agencies claim to have continuous information gathering from issuers, they tend to meet issuers only once a year. As described in Howe (1995) and S&P (1998), the rating process conducted by major rating agencies is conservative and time-consuming. Developing a rating may take months and thus rating revisions are usually delayed. The other explanation is that there exist rating biases in the bond market which cause some bonds to consistently earn higher or lower returns than similar issues. The parallel argument can be found in the stock market. In particular, Fama and French (1992) find that high book-to-market (BM) stocks outperform stocks with low book-to-market ratios, and Banz (1981) and Basu (1983) uncover higher realized returns associated with small size companies. Daniel and Titman (1997) further prove that it is the characteristics (such as BM and size) of stocks to explain the variation in stock returns.

In this paper, we propose a market-based approach to evaluate models in predicting bond ratings. We assume that the market can form an independent assessment about default risks of bonds and adequately reflect all relevant information. A relative rating strength portfolio is formed by buying under-rated bonds whose agency ratings are lower than predicted ratings from the model, and selling over-rated bonds whose agency ratings are lower than predicted ratings. Assuming that the prediction model can capture the true rating process and rating agencies will revise the ratings of mis-rated bonds eventually, under-rated bonds should be upgraded and over-rated bonds should be downgraded after portfolio formation. Since Weistein (1977), Wansley and Clauretie (1985), and Hite and Warga (1997) all indicate that there exist negative abnormal returns around downgrade announcements and positive abnormal returns for upgrades, the relative rating strength portfolio may earn abnormal profits in the long-run. Based on the long-run performance of relative rating strength portfolios, we can evaluate abilities of different models to predict bond ratings. If a statistical model can identify the relative rating strength portfolio which consistently generates more significant returns than portfolios formed based on other models, this model has more power to distinguish mis-rated bonds from correctly rated bonds, and is a better model in predicting true bond ratings. This evaluation criterion is superior to the traditional measure used in previous studies because we incorporate bond returns into the evaluation and do not assume that agency ratings are always accurate.

This paper not only examines the long-run performance of relative rating strength portfolios to evaluate models in predicting bond ratings, but also tries to identify the existence of a profitable bond trading strategy. Since previous studies of rating predictions do not provide any investment implications, this paper attempts to fill this gap by making connection between the two lines of research. Moreover, our results can be used to examine the information efficiency of the bond market with respect to accounting variables. If the relative rating strength portfolio can generate significant abnormal returns, it implies that bond prices do not fully reflect all publicly available accounting information because the predicted ratings from statistical models are all based on accounting variables.

Using 4,474 industrial bonds with Moody's ratings and 4,495 bonds with S&P ratings issued by a total of 415 firms, we perform rating predictions by four statistical models, namely, multiple discriminant analysis (MDA), multiple discriminant analysis with cross-validation procedure (MDA-C), ordered probit (Probit), and ordered probit with stepwise variable selection (Probit-S). Applying traditional evaluation measure, empirical results show that 87% of bond agency ratings can be correctly replicated by MDA, 75% by MDA-C, 80% by Probit, and 74% by Probit-S. Consistent with previous research, it suggests that statistical models can capture the essence of the rating agencies' rating process. However, since Probit does not perform better than MDA, the theoretical advantages of ordered probit models do not guarantee the superior performance, a result which is not consistent with Ederington (1985) but conforms to the findings in Kaplan and Urwitz (1979) and Wingler and Watts (1982).

As to our proposed approach, all relative rating strength portfolios implied from four models perform well in raw returns over the five-year holding horizon. However, after control rating risks, only portfolios from MDA-C can generate statistically significant returns in the long run. This further supports the notion that ordered probit model does not perform better than multiple discriminant analysis in classifying bonds. It also indicates that the evaluation criterion used in previous studies may be misleading since MDA-C performs better than MDA in our new approach but not in the traditional one. Although we find that MDA-C can effectively classify bonds better than other statistical models, the long-run abnormal return from MDA-C is about 5% in Moody's sample and 2% in S&P sample over a 5-year horizon. These returns do not seem to be economically significant, suggesting that the corporate bond market processes the accounting information used in rating prediction models with relative efficiency. On the other hand, it's also possible that the statistical models examined here do not have the ability to predict true ratings, leading to insignificant returns for our relative rating strength portfolios.

When only investment-grade bonds are examined, none of our models can generate significant abnormal portfolio returns in the long run. In other words, the long-run abnormal return from MDA-C is mainly attributed to mis-rated bonds in the below-investment grade. This implies either Moody's and S&P ratings are more accurate and timely for investment-grade bonds but not for speculative bonds or investment-grade bond market is more efficient than high yield bond markets.

We also apply our trading strategy to stocks by buying stocks whose corresponding bonds are under-rated and selling stocks whose corresponding bonds are over-rated. Surprisingly, stock portfolios from all models experience negative returns over the long run. This result, combined with the positive bond portfolio returns, means that stock and bond returns are negatively correlated and is inconsistent with the evidence in Kwan (1996) which finds a contemporaneous and negative relation between stock returns and bond yield changes. One possible explanation is that firm-specific information affects the variance, rather than the mean, of the firm value. In particular, we can treat stocks as a call option on the firm value, and bonds as the difference of risk-free bonds and put option. When the new information increases the variance of firm value, both call and put options become more valuable, and thus stock value increases and bond value decreases. Consequently, stock returns and bond returns tend to be negatively correlated. However, since we only have a small number of stocks in the portfolio for each year and stock returns are very volatile, we can't exclude the possibility that our results are due to the extreme performance of some stocks.

The rest of the paper is organized as follows. Section II reviews the models used in previous research to predict bond ratings. Section III describes the data the methodology employed in this paper. Section IV presents empirical results and Section V concludes.

II. Literature Review

Numerous studies have spent a lot of efforts to predict bond ratings by using a limited number of financial ratios through statistical models. Four models have been proposed and examined in the literature, namely, ordinary least squares, multiple discriminant analysis, ordered probit, and multinomial logit models. This section reviews these papers and introduce the benefits and restrictions associated with each model.

A. Ordinary Least Squares Model

Horrigan (1966) pioneers the bond classification research by using simple ordinary least square (OLS) regressions. He uses bond ratings as the dependent variables and assigns consecutive integers to represent the different bond ratings: 9 for AAA, 8 for AA, …, and 1 for C. With the usage of five financial ratios and one dummy variable for the subordination status, he finds that 58% of Moody's ratings and 52% of S&P ratings in his sample can be correctly predicted in advance by the OLS model. West (1970) follows Horrigan's approach, but uses different variables, all in the logarithm, to reflect the variables employed in Fisher (1959) who investigates the determinants of risk premiums on corporate bonds. However, the improvement is only marginal, 61% of bond ratings are correctly predicted in his ex-ante sample. These two early papers suggest that more than half of the bond ratings can be replicated by simple OLS model with only a few public available financial ratios.

However, there are two drawbacks associated with OLS model. First, as pointed out in Kaplan and Urwitz (1979), the dependent variables in an OLS model are defined to have an interval scale, which assumes that the risk differential between AA and A bonds is the same as that between BB and B bonds. This assumption is clearly not valid in the nature of bond ratings. Second, Mckelvey and Zavoina (1975) illustrate that the error term of OLS regression does not have a zero mean or constant variance when the dependent variable is ordinal. This indicates that, for bond rating classifications, the error term is not normally distributed and thus OLS regression assumption is violated. Pogue and Soldofsky (1969) attempt to avoid these problems by employing the linear dichotomous probability model which classifies bonds of two ratings at a time. Although they cleverly sidestep the problems in the OLS prediction model and make use of the simplicity of OLS regressions, their models do not account for all the information in bond ratings. This is because when only two ratings are considered in the regression, the whole bond rating structure and information in other ratings are ignored. Besides, they include only 10 bonds in each rating class for each regression, making their results not trustworthy.

B. Multiple Discriminant Analysis Model

As an alternative approach to predict bond ratings, multiple discriminant analysis (MDA) has been extensively applied to business and financial problems recently. MDA is a statistical technique to classify observations into different groups by maximizing the ratio of between-group variance to within-group variance. Unfortunately, by treating separate groups differently but not giving any assumed relationship among groups, MDA is not theoretically appropriate in bond classifications since it does not consider the ordinal nature of bond ratings. Moreover, MDA requires a strong assumption that independent variables in the model have to follow a multivariate normal distribution. Eisenbeis (1977), Pinches and Trieschmann (1977), Pinches (1978), and Altman et al. (1981) all argue that this assumption is clearly violated for bond rating classifications because financial ratios which earlier studies have ever considered do not have the univariate normal distribution, not even mention multivariate normal. The other problem of using the MDA model is the choice between linear and quadratic models. Linear MDA restricts variance-covariance matrices (or dispersion matrices) within different groups to be equal while quadratic MDA does not. Because Pinches and Mingo (1975) and Pinches (1978) provide the evidence that dispersion matrices are unequal, quadratic MDA should be the best choice for bond rating predictions. However, Lachenbruch et al. (1974) find that when the non-normality is present, quadratic MDA performs worse than linear MDA even dispersion matrices are unequal. Therefore, though an MDA model avoids the problems of interval scales and non-normal error terms in an OLS model, it brings new problems to bond rating predictions.

Even though the MDA model has many problems, it is the most popular model to predict bond ratings. Pinches and Mingo (1973) first apply the MDA model to bond rating classifications, and demonstrate that 65% and 56% of bonds in their holdout and ex-ante samples, respectively, can be correctly classified. Recognizing the fact that dispersion matrices are unequal for different rating groups and applying the model to senior and subordinated bonds separately, they obtain the similar prediction rate in their later study (1975). A lot of subsequent studies then follow the same way to predict ratings by either including more variables to improve the model[1] or applying the

---

[1] By including more variables, Belkaoui (1980, 1983) slightly increase the accuracy rate in predicting industrial bonds. In addition, Martin and Henderson (1983) and Perry et al. (1985) apply the rank transformation on independent variables to mitigate the non-normality problem and find some improvement.

model on different debt markets[2], as well as on bond rating changes.[3] The accuracy rate of rating prediction for these papers varies from 66% to 70% for industrial bonds, 77% to 91% for utility and financial bonds, and 75% to 80% for rating changes. This suggests that even the MDA model is not theoretically appropriate in predicting bond ratings, empirical results do not show much disadvantage against it. The fairly high accuracy rates in classifying corporate bonds explains the popularity of the MDA model in previous research.

C. Ordered Probit and Multinomial Logit Models

The ordered probit model, introduced by Mckelvey and Zavoina (1975), is designed to extend the application of dichotomous probit model to problems with the ordinal nature. Though ordered probit is a variant of the OLS regression model, it does not have the undesirable features involved in OLS regressions to predict bond ratings. Specifically, the relationship between bond ratings and independent variables in the ordered probit model is not linear any more but a normal cumulative probability distribution. With the estimation of breakpoints, the model is flexible in choosing the interval between two groups to best fit the data. Therefore, the OLS problem of fixed interval between two adjoining rating classes can be avoided in the model. Moreover, since the model assumes that dependent variables are ordinal, it makes use of the information of the bond rating structure and has an theoretical advantage over the MDA model in predicting bond ratings.

In spite of its suitable application to bond classifications, empirical results do not consistently support the idea that ordered probit dominates OLS and MDA in each application. Kaplan and Urwitz (1979) examine the performance of the ordered probit model in classifying industrial bonds and find that in the best situation 69% of bond ratings can be correctly predicted. However, when they compare ordered probit with OLS by using Horrigan's variables, the OLS model performs slightly better than ordered probit. This implies that the aspect of flexible intervals among rating groups for the ordered probit model is not so important as previous papers suggest. In fact, similar arguments can be found in Altman et al. (1981) which shows that OLS, MDA, probit, and logit models perform approximately the same in classifying banks. Wingler and Watts (1982) also document that the accuracy rate for ordered probit is lower than that of MDA in the prediction of bond rating changes. The evidence suggest that, as Wingler and Watts

[2] The MDA model has been applied on several debt markets, including electric utility bonds (Altman and Katz (1976)), financial bonds (Peavy (1980)), municipal bonds (see Loviscek and Crowley (1990) for a review), and commercial papers (Peavy and Edgar (1983, 1984), Perry and Cronan (1986)).

[3] Numerous studies also employ MDA models to predict rating changes in different sectors (Pinches et al. (1978), Bhandari et al. (1979, 1983), McAdams (1980), and Peavy (1984)).

argue, the more appealing theoretical properties of the ordered probit model do not guarantee the better prediction power. On the contrary, Ederington (1985) compares four models (OLS, MDA, probit, and logit) using only four variables in predicting industrial bonds and shows that ordered probit and multinomial logit perform better than OLS and MDA. These inconclusive results indicate that more comparisons among models need to be done to further understand their performance in predicting bond ratings. It also implies that evaluation measure used in previous research may be misleading and thus reach different conclusions in different studies.

The last model in the bond classification literature is the multinomial logit model. Similar to MDA, multinomial logit is an unordered model which does not consider the fact that bond ratings are ordinal. However, it allows the coefficients of independent variables vary across different groups, a flexibility that ordered probit does not possess. This feature may be important in bond rating predictions since bonds with different ratings have different characteristics and some financial ratios may be weighted more in AAA bonds than in CCC bonds. Unfortunately, this also brings the complexity of estimating parameters in the model. Therefore, except Ederington (1985), no previous studies have even tried to apply multinomial logit to the bond classification. When comparing the performance between ordered probit and multinomial logit, Ederington (1985) does not show the dominance of any one of these two models: multinomial logit performs better in estimation sample, but ordered probit excels in the holdout sample. He concludes that there is a trade-off of the simplicity and efficiency of ordered probit versus the flexibility of unordered logit.

D. Any Investment Implication from Previous Research ?

Table 1 summarizes all variables used to predict industrial bonds in previous studies, including Howe (1995) and S&P (1998) which introduce the credit rating process in major rating agencies. The prediction results documented in earlier research combined with table 1 indicate that statistical models can classify bonds fairly well by using only a limited number of accounting variables. As argued by Pinches and Mingo (1973, 1975), Ederington and Yawitz (1987) and many others, most of earlier studies can correctly predict two-thirds of bond ratings by applying different models, suggesting that statistical techniques do have captured the essence of the rating process.[4]

However, previous studies do not intend to provide any investment implication to bond

[4] However, Moody's and S&P keep claiming that their rating process can not be replicated by some statistical modes. For example, on its web site (www.moodys.com/moodys/mdyappr.htm), Moody's writes " ... we believe that any attempt to reduce credit rating to a formulaic methodology would be misleading and would lead to serious mistakes."

investors. Whether or not investors can make profits by predicting bond ratings and trading misclassified bonds in advance is still an open question. More importantly, all these studies try to replicate agency ratings, rather than true ratings, and fail to consider the fact that bond ratings supplied by rating agencies may not be accurate in a timely fashion. In fact, Weinstein (1977), Wansley and Clauretie (1985), and Hite and Warga (1997) all prove that the bond market starts to react to rating changes at least 6 months before the announcements. If investors can form an independent assessment about true default risks, they can forecast true bond ratings beforehand by incorporating market information into prediction models and trade issues which deviate from true ratings. Assuming that rating agencies will revise ratings of these mis-rated bonds later on, the trading strategy based on the true rating prediction models should generate abnormal profits.

III. Data and Methodology

A. Bond Data and Sample Selection

The corporate bond data are from the Fixed Income Database of University of Houston. This database includes the month-end data of bonds which compose the Lehman Brother Bond Indices. The advantage of using this database is that it differentiates trader quotes from matrix quotes for bond prices. Matrix prices are solely determined by other bonds with the same rating or by adding a fixed spread over Treasury bonds for bonds which are not actually traded in the market. Since bond dealers do not commit to trade these matrix prices, using other popular bond database such as Merrill Lynch Securities Pricing, which does not make any distinction between trader and matrix prices, will seriously affect our analysis. Both Warga (1991) and Warga and Welch (1993) argue that using matrix bond prices can be misleading and may bias return calculations.

Our initial data cover the period from January 1973 to March 1997. We restrict our sample to industrial bonds only, since we can obtain more bonds from this sector. Putable bonds are eliminated from the sample because investors of putable bonds can sell bonds back to the issuer and have more protections when the default occurs, which in turn may bias our rating predictions. Similarly, for each firm, we examine the senior bonds only and discard other bonds.[5] Generally, subordinated bonds are rated up to one letter rating lower than senior bonds. Pinches and Mingo (1973, 1975) find that in their sample subordinated bond are rated Baa or below, while non-subordinated bonds are rated Baa or higher. Thus, pooling subordinated and non-

[5] We choose bonds with subordination code of 10 (senior), 67 (senior notes), or 68 (senior debenture).

subordinated bonds together to make classification will reduce the prediction accuracy rate.[6] By removing subordinated bonds, we can make sure that our rating prediction model is not affected by the subordination status of bonds. Moreover, to guarantee that our results are not the manifestation of some tiny bonds, we delete bonds with outstanding amount less than $10 million at the time when ratings are predicted. Besides, only straight bonds are retained in the sample. This will eliminate all convertible, asset backed, and mortgage back bonds, as well as collateral mortgage obligations. Apparently, we do not remove callable bonds and bonds with sinking fund provisions from our sample. The reason is that we will lose more than 80% and 60% of bonds if we get rid of bonds with call options and sinking fund provisions, respectively.

To make sure that we obtain all necessary information we need to predict ratings, our models include variables used in previous studies for each category in table 1, and these variables are shown in the appendix. Some market-related variables, such as recent stock returns, market value of total assets, market capitalization of equity (Size), and book-to-market ratio (BM), are added in the model to fully incorporate the market information. On each April 30 from 1973 to 1996, we match the bond data with Compustat tape and allow a four-month lag to collect the financial reports. Data of common stock price and number of shares are from the last trading day of previous December in CRSP tape. To ensure that bond price is available to the investors, we require bonds to have trader prices on April 30 when bonds are used to predict ratings. Since we are interested in the long-run performance of misclassified bonds implied by our models, bonds are further restricted to have maturity of 5 years or longer on each April 30 when bond issues enter the sample. Finally, all bonds included in our analysis must have either Moody's or S&P ratings. Bond data on April, 1973 are removed from the sample because only four bonds meet our criteria. The final sample contains 4,474 bonds with Moody's ratings and 4,495 bonds with S&P ratings issued by 415 firms[7], and the distribution of bond ratings in the sample is reported in table 2.

Since Lehman Brothers was not active in the high-yield bond market until 1992, our sample tilts toward investment-grade issues. Less than 18% of bonds come from speculative issues. The single rating with the most bond issues is A which composes 41% and 39% of bonds for Moody's and S&P samples, respectively. The sample size varies over time, increasing from

[6] Pinches and Mingo (1973) adopt a 0-1 dummy variable to represent the subordination status of bonds. Since a 0-1 dummy variable is not normally distributed, the MDA assumption that independent variables follow the multivariate normal distribution is also violated.

[7] We do not restrict that one firm can have only one bond on each April 30 in our sample. This is because we want to make use of all available bonds to derive trading strategies. Besides, Moody's and S&P may occasionally mis-rate one of many bonds issued by the same firm. Confining to one bond for each firm, we may lose the opportunity to pick up arbitrage profits.

late 70s up until late 80s, after periods of smaller numbers of bonds in mid 70s. This increase in the number of bonds should be related to the hot market of junk bonds and the high bond yield period in 80s.

B. Summary Statistics

The summary statistics of bonds in the sample are presented in table 3. Except years, the numbers in the table are the time-series average of mean values of variables for each bond rating. Specifically, for each April 30 from 1974 to 1996, we compute the mean values of different variables for bonds meeting our selection criteria. Then we calculate the average of mean values in each rating across time. Since we have only relatively small number of bonds rated Caa/CCC and below, the average values for these ratings are not meaningful. The results for Moody's and S&P are shown in panel A and B, respectively.

As we expected, the higher the bond rating, the lower the bond yield. The amount outstanding and time-to-maturity also generally increase as the rating moves up. In addition, we check the data of stocks whose corresponding bonds are included in the sample, and get the Size and BM quintile breakpoints, obtained from all NYSE stocks on each April. Clearly, as firms are rated lower, the Sizes decline and BM ratios increase. Since our sample concentrates on investment-grade bonds, the size rankings for Baa/BBB and above indicate that most firms included in this study are very large. All these results suggest that larger firms tend to issue bonds with higher ratings, longer maturity, and larger amounts outstanding. Interestingly, from rating Aaa/AAA down to Ba/BB, the prior six-month buy-and-hold returns monotonically increase, but firms with rating B bonds earn less return than Ba/BB, suggesting that higher default risks do not give firms higher returns. This finding is consistent with the conclusion in Dichev (1998). When we include only the longest maturity bond for each firm, the results (not reported here) are very similar.

C. Rating Predictions and Return Calculations

Initially, we select two models, MDA[8] and ordered probit, to predict bond ratings. The choice of MDA is its popularity among previous studies and thus can facilitate the comparison of our models to others. However, MDA requires the multivariate normal distribution for independent variables, and does not take the ordinal nature of ratings into account. On the contrary, ordered probit makes use of the information that ratings are ordinal and does not need

---

[8] To be simple, we use linear, rather than quadratic, MDA to predict ratings, even though dispersion matrices are unequal. Besides, the relative few bonds in rating Ba/BB and below for some years prevent us

independent variables to be multivariate normal, which turns to be the theoretical advantages of ordered probit over MDA. By conducting MDA and ordered probit on the same sample, we can empirically test whether ordered probit perform better in predicting bond ratings. Moreover, since MDA tends to underestimate the actual error rate, Lachenbrush holdout procedure (also called cross-validation or Jackknife method) is conducted to check the robustness of MDA results. Lachenbrush method is performed by omitting one observation each time, calculating the classification rule based on remaining observations, and classifying the omitted observation. Repeat the above step until all observations are classified. As pointed out in Pinches (1980), this method can generate unbiased estimates of error rates, and is reasonably robust to extreme numbers of variables and observations. Similarly, to avoid that large number of independent variables seriously affects the probit regressions, we carry on stepwise variable selection procedure to make sure that we include in the ordered probit model the most significant variables only. Stepwise variable selection procedure is a statistical technique which picks up a subset of variables to produce a good classification among groups. A variable will be entered into or removed from the model depending on its contribution to the model. Therefore, we employ four models to predict bond ratings, namely, MDA, MDA with cross-validation (MDA-C), ordered probit, and ordered probit with stepwise variable selection (Probit-S).

On each April 30, we perform the letter rating prediction for all available bonds with Moody's or S&P ratings, by applying four statistical models mentioned above[9]. For any bond with the predicted rating implied by the model higher than rating agency's rating, we will classify it as a under-rated bond since the current rating is worse than it's supposed to be. Similarly, a bond with predicted rating lower than the rating provided by rating agencies is grouped as an over-rated bond. We then form a zero-investment portfolio by buying under-rated bonds and selling over-rated bonds, and hold for 5 years. This zero-investment portfolio is similar to the relative strength portfolio introduced in Jegadeesh and Titman (1993), so we call it the relative rating strength portfolio. Annual buy-and-hold returns for individual bonds are computed by compounding monthly returns[10] from May up to next April, or the final month in the data, whichever is earlier. Equal weighted portfolio returns are then calculated with annual portfolio

to use quadratic MDA. We also use proportional prior probabilities to get the results for MDA.

[9] There are two reasons that we don't perform refined rating prediction on all bonds in the sample. First, Lehman Brothers was not active in the high-yield bond market before 1992 and did not track all the junk bond data, so the refined ratings for below-investment are not trustworthy before 1995. Secondly, since we have only relatively small number of below-investment bonds, classifying these bonds into more refined groups will burden our model, and may seriously affect our results.

[10] The monthly return in the database is computed based on the bond full price which is the flat price (or quoted price) plus accrued interests.

rebalancing. If the number of bonds in either under- or over-rated group drops to zero in any year over the 5 year holding horizon, the return calculation stops. To track the long-run portfolio performance, we obtain the multi-year long-run returns by compounding the portfolio annual returns. Then the returns of interests are the differences of long-run compounded returns between the under- and over-rated portfolios.

Since matrix prices of bonds are not prices actually traded and thus may bias the results, we only report the results based on trader prices.[11] To be qualified for trader price analysis, bonds must at least have trader quotes at the beginning and ending of any one year holding horizon. In particular, if a bond has a trader price on April 30 this year only, but not next year, it will be eliminated from the return calculation for both this year and next year. This is to guarantee that the returns based on these bond prices can be realized by investors and that our relative rating strength strategy is implementable. To avoid the survivorship bias in the trader price analysis, we calculate the return for each bond up until its last trader price over the five-year holding horizon. For example, if a bond has trader prices on April 30 and November 30 but does not have any trader quote over the next 53 months (or is delisted on December 1), its 7 month buy-and-hold return is included in the first year portfolio return. However, it will be removed from the portfolio for the remaining years.

Repeating all the above classifications and return calculations for each year, we generate a time series of long-run returns in different investment horizons (one to five years) for relative rating strength portfolios. Assuming equal weight for each year, we calculate mean returns for different investment horizons and *t*-test is then performed to detect the significance of portfolio returns. The magnitude and significance level of returns of relative rating strength portfolios can thus be employed to evaluate the performance of different statistical models in predicting bond ratings. If a model can generate remarkable returns in terms of magnitude and significance level, it implies that this model can detect bonds whose ratings deviate from true ratings, and is more powerful in predicting bond ratings.

Since different statistical models may generate different numbers of misclassified bonds across the sample period, we also calculate the number-weighted average of returns by multiplying the number of misclassified bonds by the relative rating strength portfolio return, summing across years, and dividing it by total number of bonds in the sample period. The purpose of computing number-weighted average is to avoid that an extreme mean return generated by small number of bonds in one year bias our results.

We also compute abnormal returns of relative rating strength portfolios by carefully controlling the bond risks. Specifically, we calculate the market buy-and-hold returns matched by the rating and holding horizon for each misclassified bond and use them as benchmark returns. Similar to calculations of portfolio returns, annual market returns and long-run compounded returns are computed. For each of under- and over-rated portfolios, abnormal compounded returns are then obtained by subtracting the long-run compounded returns of market index from those of the portfolio. Again the returns of interests are the differences of abnormal compounded returns between the under- and over-rated portfolios. For the market indices, we use Lehman Brothers corporate bond indices which include all non-convertible bonds with at least one year to maturity and an outstanding amount of $50 million. These indices are categorized by different ratings and sectors, but not jointly. For each rating or sector, there are two indices, intermediate index which is made of bonds with maturities of up to ten years, and long-term index which includes bonds with maturities of 10 years or longer. Since we believe that default risks are more important than sectors to bond prices, we choose to control ratings and maturities, rather than sectors and maturities. The calculation of benchmark returns follows the portfolio return calculation. That is, if a bond is removed from the return calculation because either it does not have trader price or it is delisted, the market return calculation also stops. Moreover, if the bond is downgraded or upgraded to another rating, or its maturity is reducing from long-term to intermediate-term, the benchmark return is carefully matched at the same time. If the benchmark return is not available in the first holding horizon year, the bond is completely eliminated from the adjusted return analysis.[12] Since Lehman Brothers bond indices employ Moody's ratings of bonds first, and choose S&P ratings only if Moody's is not available, a potential problem arise when bond ratings provided by S&P are used to match market index returns. To attack this issue, we use corresponding Moody's ratings of bonds to match the market index for S&P sample. If a bond in the S&P sample does not have available Moody's rating, it will be eliminated from the sample.

IV. Empirical Results

A. Traditional Evaluation Measure for Models to Predict Ratings

To perform the rating prediction and its evaluation, past studies usually employ a two-step procedure. The model is first estimated in the training sample, and then the estimated model

---

[11] While not shown here, the results based on matrix prices are similar to those based on trader prices.

[12] Since Lehman Brothers was not an active participant in the junk bond market until 1992, the starting dates for bond market indices of rating Ba to Caa and rating Ca to D are July 1983 and January 1993, respectively. Any misclassified bond with rating Ba and below before 1983 or rating CC and below before 1993 is deleted because we can't find its abnormal returns.

is applied to the holdout sample composed of bonds not included in the training sample. The accuracy rate (or hit rate), the percentage of bonds with model ratings equal to agency ratings, of the holdout sample is thus used as the evaluation criterion. Unlike earlier papers, this study does not separate bonds into training and holdout samples. The reason is that we want to use all possible information to detect the long-run performance of misclassified bonds. As pointed out by Johnson and Wichern (1992, p515), all of the data must be used to construct the classification function to avoid that some valuable information is lost. Besides, we want to obtain as many misclassified bonds as possible to keep our results away from small sample problems.

Table 4 reports the hit rates among different models for Moody's and S&P ratings. The average hit rate across years of MDA is 87.53% for Moody's ratings and 86.58% for S&P ratings which is higher than that in previous studies about industrial bond rating predictions. The only exceptions are Perry et al. (1984, 1985) which examine different industries separately and obtain accuracy rates higher than 85% for some industries. However, as they examine the whole sample, the prediction errors significantly increase and accuracy rate is lower than 70%. With a hit rate of 75.98% in Moody's sample and 73.77% in S&P sample, MDA with cross-validation procedure (MDA-C) also performs well and dominates similar models in earlier research.[13] Since MDA-C error rates are proved to be unbiased and relatively robust, our results indicate that at least three-fourths of the rating process of Moody's and S&P can be replicated by our models. This outstanding performance is probably attributed to the fact that we include more variables in our models and exclude the subordinated bonds from the sample. Besides, we perform the rating prediction in a year-by-year basis, rather than pool many years together in the training sample as previous studies do, and can make good use of the available information.

Although ordered probit has theoretical advantage in predicting ratings, surprisingly, it does not provide a higher hit rate than that of MDA in either Moody's or S&P sample. It is possible that highly correlated variables are included in the probit model and bias prediction results. However, ordered probit with the stepwise variable selection procedure (Probit-S), which chooses the most important variables for the model, works worse than the ordered probit model, and even than MDA-C in S&P sample, in terms of hit rates. It implies either that the theoretical advantages of the ordered probit model are not important as we thought or that MDA has more power to classify bonds across large range of rating classes. This is not consistent with the conclusion in Ederington (1985) that ordered probit dominates MDA in predicting ratings, but

[13] Pogue and Soldofsky (1969) find that 80% of bonds can be correctly predicted in the holdout sample. However, their holdout sample includes only 20 bonds and may have serious sample selection biases. For a review of model performance in previous research, please see Altman et al. (1981, p213) and Ederington

supports the results in Wingler and Watts (1982) which examine the electric utility bond rating changes.

B. Market-Based Assessment for Models to Predict Ratings

When we employ traditional measure to evaluate the model performance of rating predictions, we assume that ratings provided by rating agencies are correct in a timely fashion. However, since the rating process by Moody's or S&P is conservative and long, it's very likely that agency ratings do not capture the true ratings for some periods of time. Allowing for the possibility that ratings are not accurate for each time period, we design alternative measure to evaluate the model performance, by examining the behavior of relative rating strength portfolios. This new evaluation criterion is to assume that market can have its own assessed ratings, which sometimes may be different from ratings supplied by rating agencies, about the credit risks of issuers. If a model can persistently produce significant portfolio returns over the long-run, it suggests that not only this model works well to identify true bond ratings but also bond market is not efficient.

Table 5 shows the long-run performance of relative rating strength portfolios, formed by buying under-rated bonds and selling over-rated bonds.[14] For raw returns, all models can generate significant returns in the long-run for both Moody's and S&P samples. In particular, over five post-formation years, relative rating strength portfolios from Moody's sample can earn an average return of 11.89% for MDA[15], 9.9% for MDA-C, 9.57% for Probit, and 8.35% for Probit-S, all of which are significant within 1% level. Similar results are observed in the S&P sample. Since number-weighted average return is similar to equal-weighted average, our results are not due to a small subset of misclassified bonds. However, after control the bond market index returns, MDA[16], Probit and Probit-S cannot consistently yield significant abnormal returns in the long-run. These imply that the great long-run performance of raw returns is just the manifestation of rating difference between under- and over-rated portfolios. While not shown here, our analysis

---

and Yawitz (1987, chapter 23, exhibit 22).

[14] For MDA and ordered probit models, since there are only few misclassified bonds before 1977, we calculate the portfolio return starting from 1978 for both Moody's and S&P samples.

[15] For Moody's sample, the relative rating strength portfolio identified by MDA and formed on April 30, 1984 has extreme long-run returns. If we delete the portfolio formed in that year, number-weighted and equal-weighted average return on the 4th and 5th post-formation years will decrease to 10% and 9%, respectively. The portfolio from MDA-C in S&P sample has the same situation in 1984, and the average return will down to 10% if year 1984 is ignored.

[16] Again, in Moody's sample, the returns in year 4 and 5 for MDA are due to extraordinary returns in one year, 1983. If we delete portfolios formed in 1983, the average return is down to 2.5%. However, the significance level does not improve. In table 6, similar situations can be found in returns for MDA in Moody's sample.

indicates that the average ratings in under-rated portfolios tend to be lower than those in over-rated ones. Surprisingly, MDA-C performs well even after controlling rating risks. The returns are significant from year 2 to year 5 for Moody's sample, and year 1, 3, 5 for S&P sample.

Some implications can be derived from table 5. First, MDA-C works the best among four models to predict industrial bond ratings. This indicates that ordered probit models are not superior to MDA models in classifying bonds because MDA-C can capture bonds whose true ratings deviate from agency ratings while Probit and Probit-S can't. Second, since MDA-C can identify relative rating strength portfolios which statistically outperform the market indices over the long-run but MDA can't, it suggests that predicted ratings from holdout sample procedure is better. Using traditional evaluation criterion, we observe the opposite finding in table 4 (the hit rate is about 87% for MDA and 75% for MDA-C). Because the return results here are based on actual traded data after ratings are predicted, our proposed evaluation criterion includes post market information and should be superior to the traditional measure. Therefore, the different conclusions reached from table 4 & 5 reveal the fact that adopting the traditional measure in evaluating statistical models to predict ratings can be misleading.

Although we detect a profitable trading strategy from the corporate bond market by MDA-C model, the abnormal returns are not economically significant, with about 5% for Moody's sample and 2% for S&P sample over a 5 year horizon. It seems that the corporate bond market efficiently processes the relevant accounting information since the statistical models based on these information cannot generate large abnormal returns. However, we cannot rule the possibility that the statistical models we consider here do not have the ability to predict true ratings, leading to insignificant returns for our relative rating strength portfolios.

A related issue is the split rating, a situation that a bond has different ratings from Moody's and S&P, in our sample. From year 3 to year 5, relative rating strength portfolios using Moody's ratings seem to outperform those using S&P ratings. The return difference is 2.51% (4.80%-2.29%) for equal-weighted averages and 4.16% (5.61%-1.45%) for number-weighted averages over 5 years. However, using *t*-test (not shown here) which compares the average returns between Moody's and S&P samples, we don't find any significant difference in year 3 and 4. This is consistent with the findings in Ederington (1986) that there is no systematic difference between Moody's and S&P ratings and in Ederington et al. (1987) that Moody's and S&P ratings are treated equally by the market.

C. Sources of Significant Long-Run Returns

To check our results, we delete below-investment bonds when we form relative rating strength portfolios, and report the portfolio long-run performance in table 6. The magnitude and significance of portfolio returns are reduced in table 6, compared to those in table 5. Consistent with the result using all misclassified bonds, there is no evidence that ordered probit models perform better than MDA models in classifying bonds. However, when portfolio returns are adjusted to market index returns, no of our models exhibits significant returns over 5 post-formation years. Even the long-run returns from MDA-C are significant only in year 4 and 5 and for Moody's sample. This implies that the statistically significant returns from MDA-C in table 5 are attributed to misclassified bonds with below-investment grade.

We also restrict our sample to investment grade bonds only. The results are presented in table 7. The difference between table 6 and 7 is that only investment grade bonds are used to predict ratings in the latter. The purpose of doing so is that we may find any difference between portfolios from letter rating predictions and refined rating predictions. Consistent with previous tables, after we control the market index returns, no model can persistently generate significant portfolio returns in the long run. As a result, the difference between letter and refined rating prediction (panel A vs. B and C vs. D) is not statistically significant. The results from table 6 and 7 suggest that relative rating strength portfolios formed by investment-grade bonds can't consistently make significant abnormal returns. It indicates that Moody's and S&P ratings may be more accurate and timely for investment-grade bonds, but not for speculative bonds. However, it also could be the case that investment-grade bond market is more efficient than high yield bond markets. One possible explanation is that speculative bond market is less liquid (for example, insurance firms and pension funds are prohibited to invest speculative bonds), and its slower information flow prevents bond prices to fully reflect all available information.

D. Any Implication to Stocks ?

Kwan (1996) documents that individual stock returns and bond yield changes are negatively and contemporaneously correlated and are driven by firm-specific information. This implies that stock returns and bond returns are positively correlated in the firm level. It also suggests that relative rating strength portfolios formed by misclassified bonds can be implemented for corresponding stocks. If returns of individual stocks and bonds tend to move in the same direction, firms whose bonds are under-rated will have stocks with higher future returns, and stocks issued by firms with over-rated bonds will have lower returns in the future.

To check this issue, we first perform the letter rating prediction and find the under- and over-rated bonds. The stocks whose corresponding bonds are under-rated will also be treated as under-rated stocks, and vice versa for stocks with over-rated bonds. If a firm has more than one bond misclassified, we only include its stock once in the portfolio in each formation year to avoid double counting.[17] To form relative rating strength portfolios in stocks, on each April 30, we buy under-rated stocks and sell over-rated stocks and hold for 5 years. Return calculations and significance tests are performed as in previous tables. To control the risk factors in stocks, we compute the adjusted returns by subtracting returns of the corresponding Size/BM control portfolio from raw returns of the stock portfolio. The result is illustrated in table 8. Surprisingly, relative rating strength portfolios in stocks exhibit negative returns, most of which are significant within 10%, from year 1 to year 5 across different models. Controlling Size and BM factors reduces the return magnitude and significance level only a little bit. Interestingly, stock portfolios from probit models have more negative returns than those from MDA models. Probit-S can even generate significant stock portfolio returns from year 1 through year 5. In table 9, similar results are obtained when stocks whose corresponding bonds are below-investment grade are eliminated from the stock portfolios.

The negative returns in stock portfolios and positive returns in bond portfolios (both in raw returns) means that stock returns and bond returns are negatively correlated. This not consistent with the findings in Kwan (1996). We argue that the relationship between stocks and bonds are driven by firm-specific information affecting the variance, rather than the mean, of firm values. In particular, we can treat stocks as a call option on the firm value, and bonds as the difference of risk-free bonds and put option. When some information increases the variance of firm value, both call and put options become more valuable, and thus stock value increases and bond value decreases. Consequently, stock returns and bond returns tend to be negatively correlated. However, since we only have a small number of stocks in the portfolio for each year and the stock market is very volatile, we can't exclude the possibility that our results are due to the extreme performance of small set of stocks.

VI. Conclusion

Numerous studies have proposed different statistical models to predict bond ratings in different debt markets. These papers assume that ratings provided by Moody's or S&P are accurate in a timely fashion and attempt to replicate rating agencies' ratings, but not true ratings. However, Weistein (1977) shows that bond market has fully anticipated the rating changes 6 months before the announcements. Moreover, both Wansley and Clauretie (1985) and Hite and

[17] If a firm has under- and over-rated bonds in the same portfolio formation year, its stock is removed from the portfolio in that year.

Warga (1997) find that there exist abnormal returns during 12 months around rating change announcements. All these findings suggest that rating agencies do not adjust ratings efficiently to fully reflect the true default risks of bonds. In other words, the assumption in previous research is not valid.

Assuming that market can form an independent assessment about default risks of bonds, we design a new approach to evaluate models in predicting true bond ratings. Specifically, we construct a relative rating strength portfolio formed by buying under-rated bonds, whose predicted ratings from the statistical model are higher than agency ratings, and selling over-rated bonds, whose predicted ratings are lower than agency ratings. We hold this portfolio for 5 years after the rating prediction is performed, and examine its long-run performance. If one model can generate the portfolio with more significant returns than the portfolio from other models, this model has more power to identify bonds whose agency ratings deviate from true ratings, and is a better model in predicting bond ratings. This evaluation criterion is superior to that used in previous research because we include post market returns into the evaluation and not assume that ratings from Moody's or S&P are always correct. Moreover, if there does exist a relative rating strength portfolio which can generate significant returns in the long-run, investors can make arbitrage profits in the corporate bond market by implementing the trading strategy. Therefore, this paper has two purposes – examine the performance of the alternative evaluation measure of statistical models to predict bond ratings, and investigate the existence of profitable bond trading strategies generated from statistical models. This paper also tries to make the connection between bond rating prediction and bond investment implications.

Using 4,474 industrial bonds with Moody's ratings and 4,495 bonds with S&P ratings issued by 415 firms, we perform rating predictions by four statistical models, namely, multiple discriminant analysis (MDA), multiple discriminant analysis with cross-validation procedure (MDA-C), ordered probit (Probit), and ordered probit with stepwise variable selection (Probit-S). Applying traditional evaluation measure, our empirical results show that 87% of agencies' ratings can be correctly replicated by MDA, 75% by MDA-C, 80% by Probit, and 74% by Probit-S. Consistent with previous research, our statistical models fairly capture the essence of the rating process of rating agencies. However, since Probit and Probit-S do not perform better than MDA or MDA-C, it implies that theoretical advantages of ordered probit models do not guarantee the superior performance, a result not consistent with Ederington (1985) but conformed to the findings in Kaplan and Urwitz (1979) and Wingler and Watts (1982).

With the alternative approach, we find that, over the 5 year holding horizon, only

MDA-C can identify the bond portfolio with significant returns after controlling rating risks. As compared with the traditional evaluation measure, the result based on our approach suggests that ordered probit model does not perform better than multiple discriminant analysis in classifying bonds. It also indicates that the evaluation criterion used in previous studies may be misleading since MDA-C performs better than MDA in our new approach but not in the traditional one. Although we detect a profitable trading strategy from MDA-C, the generated abnormal return is not economically significant, with 5% in Moody's sample and 2% in S&P sample over 5 years. Moreover, when only investment-grade bonds are examined, none of our models can yield significant abnormal returns in the long-run. Therefore, the statistically significant portfolio performance of MDA-C is mainly attributed to mis-rated bonds in the below-investment grade. This implies either that Moody's and S&P ratings are more accurate and timely for investment-grade bonds but not for speculative bonds or that investment-grade bond market is more efficient than high yield bond markets. In addition, the corporate bond market seems to process the relevant accounting information with relative efficiency since the statistical models based on these information cannot generate large abnormal returns. However, it's also likely that the statistical models examined in this paper are incapable of predicting true ratings, leading to insignificant returns for our relative rating strength portfolios.

To check whether this bond trading strategy works in stocks, we form the stock portfolio by buying stocks whose corresponding bonds are under-rated and selling stocks whose corresponding bonds are over-rated. Surprisingly, stock portfolios from all models exhibit significant negative, rather than positive, returns over the long-run. This result, combined with positive bond returns, is inconsistent with the evidence in Kwan (1996) which suggests a contemporaneous and negative relation between stock returns and bond yield changes. One possible explanation is that firm-specific information affects the variance, rather than the mean, of the firm value.

There are two caveats needed to keep in mind. First, this paper does not explicitly incorporate the industry factor into models. Previous studies suggest that rating predictions within single industry are more accurate. However, as pointed out in Ederington and Yawitz (1987), since more than two-thirds of agency ratings can be replicated correctly in earlier papers, the industry analysis may not be so important as rating agencies claim. Nevertheless, our results are subject to the model specification. Future research can include some dummy variable to proxy for different industries, or perform separate analysis on different industries, to see if there exists any difference.

Secondly, the market index returns used in this paper are taken directly from Lehman

Brother Bond Indices. Bonds with matrix prices are not excluded from index return calculations, and this could impact the index returns and bias our return results. Future research can construct market indices and compute index returns using bonds with trader prices only to resolve this issue. More bond characteristics can also be controlled if self-constructed market indices can be built. Examination of portfolio performance based on different sectors, such as electric, utility or finance, can also be used to check the robustness of our results.

Appendix
Variables Used in the Statistical Models

A. Profitability:
1. Pretax return on permanent capital = EBIT / average total assets
2. Return on asset = net income / total assets
3. Return on stock = last 12 month buy-and-hold stock return ending on previous December
4. Return on stock = last 6 month buy-and-hold stock return ending on previous December
5. E/P ratio = earnings per share / price
6. D/P ratio = dividend per share / price

B. Earnings variability:
1. ROA variability = standard deviation of recent 5 year returns on assets
2. Pretax return variability = standard deviation of recent 5 year EBIT to total assets ratios

C. Coverage:
1. Pretax interest coverage = EBT / interests
2. Pretax interest coverage including rent = (EBT + rent) / (interests + rent)
3. EBITDA interest coverage = (EBT + rent + depreciation) / interests

D. Cash flow to debt ratio:
1. Cash flow / total debt = (net income + depreciation) / (long-term debt + short-term debt)
2. Cash flow / long-term debt = (net income + depreciation) / long-term debt
3. Free cash flow / total debt = (net income + depreciation - capital expenditure) / (long-term debt + short-term debt)
4. Free cash flow / long-term debt = (net income + depreciation - capital expenditure) / long-term debt

E. Leverage:
1. Long-term debt to capitalization = long-term debt / total assets
2. Short-term debt to long-term debt
3. Deferred taxes to long-term debt

F. Firm size: (based on previous December price and number of shares in CRSP; all data but price are in millions before taking logarithm)
1. Market value of total asset = log (price * number of shares + total assets - book value of common equity)
2. Market value of common equity = log (price * number of shares)

G. Growth potential:
1. BM ratio = book value of common equity / market value of common equity
2. Sales growth in recent 3 years = log (sales in year (t–1)/ sales in year (t-3))

H. Operating efficiency:
1. Asset turnover = sales / average total assets
2. Receivables turnover = sales / average receivables

I. Liquidity:
1. Current ratio = current assets / current liabilities
2. Quick ratio = (current assets - inventory) / current liabilities

J. Other:
1. Size of bonds = log (outstanding amount of the bond in thousands on April 30)
2. Dividend to interest ratio

References

Altman, Edward I., 1989, Measuring corporate bond mortality and performance, *Journal of Finance* 44, 909-922.

Altman, Edward I., 1990, Setting the record straight on junk bonds: a review of the research on default rates and returns , *Journal of Applied Corporate Finance* , 82-95.

Altman, Edward I., 1992, Revisiting the high-yield bond market, *Financial Management* 21, 78-92.

Altman, Edward I., Robert B. Avery, Robert A. Eisenbeis, and Joseph F. Sinkey, Jr.,1981, *Application of Classification Techniques in Business, Banking and Finance* (Jai Press, Greenwich, CT).

Altman, Edward I., and Steven Katz,1976, An analysis of bond ratings in electric public utility industry, in G. Sorter and M. Schiff, ed.: *Proceedings of the Conference on Topical Research in Accounting* (New York University, New York, NY).

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Basu, Sanjoy, 1983, The relationship between earning yield, market value, and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129-156.

Belkaoui, Ahmed, 1980, Industrial bond ratings: a new look, *Financial Management* 9, 44-51.

Belkaoui, Ahmed, 1983, *Industrial bonds and the rating process* (Quorum Books, Westport, CT).

Bennett, Thomas L., Stephen F. Esser, and Christian G. Roth, 1994, Corporate credit risk and reward, *Journal of Portfolio Management* 20, 39-47.

Bhandari, Shyan B., Robert M. Soldofsky, and Warren J. Boe, 1979, Bond quality rating changes for electric utilities: A multivariate analysis , *Financial Management* 8, 74-81.

Bhandari, Shyan B., Robert M. Soldofsky, and Warren J. Boe, 1983, What's in a change in an industrial bond rating?, *Journal of Business Research* 11, 91-106.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1-33.

Dichev, Ilia D.,1998, Is the risk of bankruptcy a systematic risk?, *Journal of Finance* 53, 1131-1147.

Ederington, Louis H., 1985, Classification models and bond ratings, *The Financial Review* 20, 237-262.

Ederington, Louis H., 1986, Why split ratings occur, *Financial Management* 16, 37-47.

Ederington, Louis H., and Jess B. Yawitz, 1987, The bond rating process, in Edward I. Altman, ed.: *Handbook of Financial Markets* (John Wiley & Sons, New York, NY).

Ederington, Louis H., Jess B. Yawitz, and Brian E. Roberts, 1987, The informational content of bond ratings, *Journal of Financial Research* 10, 211-226.

Eisenbeis, Robert A., 1977, Pitfalls in the application of discriminant analysis in business, economics, and finance, *Journal of Finance* 32, 875-900.

Fama, Eugene F., and Kenneth R. French, 1992, The cross section of expected stock returns, *Journal of Finance* 47, 427-465.

Fisher, Lawrence, 1959, Determinants of Risk Premiums on Corporate Bonds, *Journal of Political Economy* 67, 217-237.

Hickman, W. Braddock, 1958, *Corporate Bond Quality and Investor Experience* (Princeton University Press, Princeton, NJ and the National Bureau of Economic Research, New York, NY).

Hite, Gailen, and Arthur Warga, 1997, The effect of bond-rating changes on bond price performance, *Financial Analysts Journal* 53, 35-51.

Holthausen, Robert W., and Richard W. Leftwich, 1986, The effect of bond rating changes on common stock prices, *Journal of Financial Economics* 17, 57-89.

Horrigan, James O., 1966, The determination of long-term credit standing with financial ratios, *Journal of Accounting Research* 4, 44-62.

Howe, Jane T., 1995, Credit analysis for corporate bonds, in Frank Fabozzi and Dessa Fabozzi, ed.: *The Handbook of Fixed Income Securities* (Irwin, Inc., Chicago, IL)

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implication for stock market efficiency, *Journal of Finance* 48, 65-91.

Johnson, Richard A., and Dean W. Wichern, 1992, *Applied Multivariate Statistical Analysis*, 3rd ed. (Prentice-Hall, Englewood Cliffs, NJ).

Kaplan, Robert S., and Gabriel Urwitz, 1979, Statistical models of bond ratings: a methodological inquiry, *Journal of Business* 52, 231-261.

Kwan, Simon H., 1996, Firm-specific information and the correlation between individual stocks and bonds, *Journal of Financial Economics* 40, 63-80.

Lachenbrush, Peter A., Cheryl Sneeringer, and Lawrence T. Revo, 1974, Robustness of the linear and quadratic discriminant function to certain types of non-normality, *Communications in Statistics* 1, 39-57.

Loviscek, Anthony L., and Frederick D. Crowley, 1990, What is in a municipal bond rating?, *The Financial Review* 25, 25-53.

Martin, Linda J., and Gleen V. Henderson, Jr., 1983, On bond ratings and pension obligations: a note, *Journal of Financial and Quantitative Analysis* 18, 463-470.

McAdams, Lloyd , 1980, How to anticipate utility bond rating changes, *Journal of Portfolio Management* 7, 56-60.

McKelvey, Richard D., and William Zavoina, 1975, A statistical model for the analysis of ordinal level dependent variables, *Journal of Mathematical Sociology* 4, 103-120.

Peavy, John W., III, 1980, The classification of bank holding company bond ratings, *Review of Business and Economic Research* 16, 18-28.

Peavy, John W., III, 1984, Forecasting industrial bond rating changes: A multivariate model, *Review of Business and Economic Research* 19, 46-56.

Peavy, John W., III, and S. Michael Edgar, 1983, A multiple discriminant analysis of BHC commercial paper ratings, *Journal of Banking and Finance* 7, 161-173.

Peavy, John W., III, and S. Michael Edgar, 1984, Rating electric utility commercial paper, *Journal of Accounting, Auditing & Finance* 8, 125-135.

Perry, Larry G., Timothy P. Cronan, 1986, A note on rank transformation discriminant analysis, *Journal of Banking and Finance* 10, 605-610.

Perry, Larry G., Timothy P. Cronan, and Gleen V. Henderson, Jr., 1985, Industry classification, ordinal data, and bond rating decision models, *Decision Sciences* 15, 14-24.

Perry, Larry G., Gleen V. Henderson, Jr., and Timothy P. Cronan, 1984, Multivariate analysis of corporate bond ratings and industry classifications, *The Journal of Financial Research* 7, 27-36.

Pinches, George E., 1978, , A multivariate analysis of industrial bond ratings and the role of subordination: Reply, *Journal of Finance* 32, 336-344.

Pinches, George E., 1980, Factors influencing classification results from multiple discriminant analysis, *Journal of Business Research* 8, 429-456.

Pinches, George E., and Kent A. Mingo, 1973, A multivariate analysis of industrial bond ratings, *Journal of Finance* 28, 1-18.

Pinches, George E., and Kent A. Mingo, 1975, The role of subordination and industrial bond ratings, *Journal of Finance* 30, 201-206.

Pinches, George E., and J. Clay Singleton, 1978, The adjustment of stock prices to bond rating changes, *Journal of Finance* 33, 29-44.

Pinches, George E., J. Clay Singleton, and Ali Jahankhani, 1978, Fixed coverage as a determinant of electric utility bond ratings, *Financial Management* 7, 45-55.

Pinches, George E., and James S. Trieschmann, 1977, Discriminant analysis, classification results, and financially distressed P-L insurers, *Journal of Risk and Insurance* 44, 289-298.

Pogue, Thomas F., and Robert M. Soldofsky, 1969, What's in a bond rating, *Journal of Financial and Quantitative Analysis* 4, 201-228.

Standard & Poor's, 1998, *Corporate Ratings Criteria* (McGraw-Hill, New York, NY).

Wansley, James W., and Terrence M. Clauretie, 1985, The impact of creditwatch placement on equity returns and bond prices, *Journal of Financial Research* 8, 31-42.

Warga, Arthur D., 1991, Corporate bond price discrepancies in the dealer and exchange markets, *Journal of Fixed Income* 1, 7-16.

Warga, Arthur D., and Ivo Welch, 1993, Bondholder losses in leveraged buyouts, *Review of Financial Studies* 6, 959-982.

West, Richard R., 1970, An alternative approach to predicting corporate bond ratings, *Journal of Accounting Research* 7, 118-127.

Weinstein, Mark I., 1977, The effect of a rating change announcement on bond price, *Journal of Financial Economics* 5, 329-350.

Wingler, Tony R., and James M. Watts, 1982, Electric utility bond rating changes: Methodological issues and evidence, *Journal of Financial Research* 5, 221-235.

Table 1
Variables Used in Previous Research to Predict Industrial Bond Ratings

AR: accounts receivables, CV: coefficient of variation, E: Equity, Int: interest, Int cov: Interest coverage, LTD: long-term debt, MV: market value, NI: net income, OI: operating income, STD: short-term debt, TA: total assets, TD: total debt, TL: total liabilities, UPSC: unfunded past service cost, WC: working capital, Capital = LTD + E, Free cash flow = cash flow - capital expenditure - change of working capital

Variables	Horrigan (66)	Pogue and Soldofsky(69)	West (70)	Pinches and Mingo (73,75)	Kaplan and Urwitz (79)	Belkaoui (80,83)	Martin and Henderson (83)	Perry,Henderson, and Cronan (84,85)	Ederington (85)	Howe (95)	S&P (98)
Leverage	E/TD	LTD/Capital	(MV of E)/ TD*	5 yr mean of (LTD/TA)	LTD/TA	• LTD/Capital • STD/Capital	• LTD/Capital • (LTD+UPSC)/ (Capital+UPSC) • [LTD+UPSC(1-0.4)]/[Capital+ UPSC(1-0.4)]	• LTD/TA • TL/TA • 5 yr mean of (LTD/E)	LTD/Capital	• LTD/Capital • LTD/ [LTD+MV(E)] • TA/TD	• LTD/Capital • TD/TD+E
Coverage		Int cov		5 yr mean of int cov	(Cash flow+ Int+Tax)/Int	Int cov	(5 yr mean of EBIT)/Int		Forecast of pretax int cov	• Pretax int cov • Pretax fixed charge cov	• Pretax int cov • EBITDA int cov
Cash flow						5 yr cash flow/ (5 yr cap.exp. +dividend)		• CV of 5 yr (cash flow/E) • CV of 5 yr [(EBIT+Dep)/TA]		Cash flow/TD	• Cash flow/TD • Free cash flow/TD
Liquidity	WC/Sales					Current ratio	Current ratio	• Quick ratio • WC/TA		• Current ratio • Quick ratio	
Profitability	OI/Sales	ROA		ROA**	ROA		ROA	E/P ratio		ROE	• OI/Sales • Pretax return on permanent capital
Earning-variability		CV of ROA	CV of 9 yr NI*					CV of 5 yr EBIT		Earnings progress	
Firm Size	TA	TA			TA	TA		5 yr mean of TA	5 yr mean of TA		
Subordination	Yes			Yes	Yes	Yes			Yes		
Size of debt			MV of bond*	Issue size*	Issue size	TD					
Others	Sales/E		Period of Solvency*	Yrs of consecutive Dividends*	Market betas	M/B ratio	• UPSC per person • UPSC/EBT • UPSC(1-0.4)/E	• Sales/TA • Sales/AR		• Intangibles • UPSC	

* Variables are in logarithm
** Variables are standardized

Table 2
The Distribution of Moody's and S&P Ratings in the Sample

This table reports the number of bonds in different ratings and years. Bonds included in the sample must meet the following criteria: 1) industrial bonds, 2) non-putable bonds, 3) senior bonds, 4) straight bonds, 5) with available financial variables listed in the appendix. Also, on each April 30 when bond ratings are predicted, only bonds with amounts outstanding larger than 10 million, time-to-maturities greater than 5 years, and with trader prices can enter into the sample. The "Avg" in the last row of the table is the average number of bonds across all years for different ratings.

	Moody's Ratings									S&P Ratings									
Year	Aaa	Aa	A	Baa	Ba	B	Caa	Ca	Total	AAA	AA	A	BBB	BB	B	CCC	CC	D	Total
74	12	18	21						51	16	15	18	2						51
75	20	20	25	4	2	2			73	20	21	22	5	4	1				73
76	22	26	27	10	2	1			88	22	24	26	12	3	1				88
77	24	28	27	13	3				95	24	30	24	13	4	1				96
78	30	65	87	27	5	1			215	28	66	85	29	9	2				219
79	33	61	89	26	5	3			217	28	68	84	26	10	4	1			221
80	36	56	81	24	4	6			207	17	62	90	28	4	7	1			209
81	32	54	102	16	9	28	1		242	15	55	119	17	7	35	1			249
82	26	38	76	11	6	29			186	7	54	77	13	5	33	1	1		191
83	14	63	91	13	14	29			224	6	73	89	12	16	33				229
84	13	53	95	19	14	32			226	6	67	78	29	14	37		2		233
85	4	60	98	42	25	31			260	4	71	75	58	20	36		1		265
86	6	45	132	38	42	28			291	6	66	91	61	33	35		1		293
87	11	40	124	44	41	44			304	8	48	104	58	29	49	2	6		304
88	9	34	126	48	23	38			278	8	35	126	45	26	28	2	1		271
89	4	22	112	46	8	20			212	4	25	108	39	15	13				204
90	5	17	76	27	4	17			146	5	17	76	26	6	15				145
91	4	18	80	32	7	13	1		155	4	18	82	30	7	12	2			155
92	3	19	83	59	10	9	2		185	3	19	94	48	7	10	2			183
93	2	21	64	65	27	18	1	1	199	2	24	71	65	23	10	1		2	198
94	3	19	76	56	21	19		1	195	5	21	71	60	23	14				194
95	4	26	73	64	22	21	1		211	4	29	70	64	28	15	1			211
96	4	18	65	68	32	26	1		214	4	21	60	72	34	20	2			213
Total	321	821	1830	752	326	415	7	2	4474	246	929	1740	812	327	411	16	12	2	4495
Avg	14	36	80	33	14	18	0	0	195	11	40	76	35	14	18	1	1	0	195

Table 3
Summary Statistics

For each year from 1974 to 1996, the mean values are first computed within each rating for each variable. The time-series average of these mean values is then calculated. "Years" means the number of years of available data used to compute the time-series average. "Amount" is the outstanding amount of bonds on each April 30, expressed in millions. "Size", in terms of millions, is the market value of equity at previous December end, and "BM" is the book-to-market ratio calculated by dividing book value at previous fiscal year by the market value of equity. "Size Rank" and "BM Rank" are based on Size/BM control portfolio quintile breakpoints obtained from all NYSE stocks. "Stock Return" is the prior six month buy-and-hold return of the corresponding stock. Panel A shows the results based on Moody's ratings, while panel B reports the results based on S&P ratings.

Ratings	Years	Amount	Yield	Maturity	Size	BM	Size Rank	BM Rank	Stock Return
Panel A: Moody's Ratings									
Aaa	23	225	9.38%	16.920	30547	0.63	5.00	2.30	7.48%
Aa	23	189	9.71%	17.154	11501	0.65	4.99	2.46	11.19%
A	23	139	9.99%	16.095	3970	0.84	4.79	3.10	12.97%
Baa	22	143	10.55%	13.589	1811	0.92	4.26	3.39	16.65%
Ba	22	136	11.60%	12.565	1165	0.89	3.57	3.04	19.42%
B	21	99	14.62%	12.705	286	1.64	2.20	3.42	15.93%
Caa	6	110	18.68%	7.237	87	2.46	1.17	4.83	46.71%
Ca	2	70	N/A	5.541	27	1.86	1.00	3.50	9.09%
Panel B: S&P Ratings									
AAA	23	223	9.30%	15.991	31823	0.63	5.00	2.30	7.75%
AA	23	201	9.74%	17.003	11189	0.68	4.96	2.57	10.48%
A	23	131	9.96%	16.498	3940	0.81	4.78	3.03	13.09%
BBB	23	170	10.45%	13.837	2009	0.96	4.37	3.44	14.91%
BB	22	115	11.70%	12.276	1010	1.04	3.61	3.08	19.14%
B	22	106	13.66%	12.770	268	1.41	2.37	3.51	16.89%
CCC	11	76	21.14%	9.772	50	2.13	1.14	4.50	13.04%
CC	6	48	20.41%	11.491	1014	0.88	1.44	3.33	-3.45%
D	1	41	N/A	7.022	68	2.23	1.00	5.00	35.23%

Table 4
Hit Rates of Various Models to Predict Bond Ratings

On each April 30 from 1974 to 1996, the bonds' ratings are predicted by four models: multiple discriminant analysis (MDA), multiple discriminant analysis with cross-validation (MDA-C), ordered probit (Probit), and order probit with stepwise variable selection (Probit-S). The variables used to predict ratings are listed in the appendix. The results based on Moody's and S&P ratings are shown in panel A and B, respectively. The "No" column reports the total number of bonds used in rating prediction in each year. The "U" column includes the number of under-rated bonds which have predicted ratings implied by the model being higher than Moody's or S&P ratings, while the "O" column contains the number of over-rated bonds whose predicted ratings are lower than Moody's or S&P ratings. "Mis" column is the number of misclassified bonds, which is the sum of numbers in "Under" and "Over" columns. "Hite rate" is the ratio of number of bonds, whose predicted ratings equal to Moody's or S&P rating, to total number of bonds in that year. The "Avg" in the last row of the table is the average number of bonds and hit rate across all years for different models.

		MDA				MDA-C				Probit				Probit-S			
Yr	No	U	O	Mis	Hit rate	U	O	Mis	Hit rate	U	O	Mis	Hit rate	U	O	Mis	Hit rate
Panel A: Moody's Ratings																	
74	51	0	0	0	100.00%	7	3	10	80.39%	0	0	0	100.00%	4	2	6	88.24%
75	73	1	1	2	97.26%	9	8	17	76.71%	1	2	3	95.89%	3	6	9	87.67%
76	88	1	2	3	96.59%	5	7	12	86.36%	1	2	3	96.59%	6	6	12	86.36%
77	95	2	3	5	94.74%	11	8	19	80.00%	1	2	3	96.84%	7	11	18	81.05%
78	215	13	20	33	84.65%	31	27	58	73.02%	24	16	40	81.40%	26	22	48	77.67%
79	217	13	12	25	88.48%	32	21	53	75.58%	25	18	43	80.18%	29	20	49	77.42%
80	207	15	13	28	86.47%	28	20	48	76.81%	23	16	39	81.16%	26	20	46	77.78%
81	242	17	17	34	85.95%	33	26	59	75.62%	33	20	53	78.10%	25	23	48	80.17%
82	186	8	3	11	94.09%	17	8	25	86.56%	11	15	26	86.02%	13	15	28	84.95%
83	224	12	24	36	83.93%	24	32	56	75.00%	12	16	28	87.50%	18	27	45	79.91%
84	226	12	20	32	85.84%	20	34	54	76.11%	20	24	44	80.53%	25	31	56	75.22%
85	260	19	26	45	82.69%	33	36	69	73.46%	36	37	73	71.92%	43	44	87	66.54%
86	291	32	18	50	82.82%	44	26	70	75.95%	28	36	64	78.01%	31	48	79	72.85%
87	304	33	37	70	76.97%	46	45	91	70.07%	44	41	85	72.04%	39	43	82	73.03%
88	278	20	25	45	83.81%	39	40	79	71.58%	44	32	76	72.66%	45	33	78	71.94%
89	212	15	12	27	87.26%	30	16	46	78.30%	27	17	44	79.25%	21	19	40	81.13%
90	146	7	4	11	92.47%	11	11	22	84.93%	11	5	16	89.04%	16	16	32	78.08%
91	155	11	8	19	87.74%	19	20	39	74.84%	14	16	30	80.65%	29	23	52	66.45%
92	185	12	7	19	89.73%	21	15	36	80.54%	15	11	26	85.95%	20	28	48	74.05%
93	199	23	12	35	82.41%	41	23	64	67.84%	28	32	60	69.85%	32	30	62	68.84%
94	195	17	12	29	85.13%	34	18	52	73.33%	23	29	52	73.33%	27	25	52	73.33%
95	211	26	17	43	79.62%	39	37	76	63.98%	29	28	57	72.99%	36	30	66	68.72%
96	214	19	14	33	84.58%	38	25	63	70.56%	38	31	69	67.76%	44	32	76	64.49%
All	4474	328	307	635	85.81%	612	506	1118	75.01%	488	446	934	79.12%	565	554	1119	74.99%
Avg	195	14	13	28	87.53%	27	22	49	75.98%	21	19	41	81.64%	25	24	49	76.34%

Table 4 - *continued*

Yr	No	MDA				MDA-C				Probit				Probit-S			
		U	O	Mis	Hit rate	U	O	Mis	Hit rate	U	O	Mis	Hit rate	U	O	Mis	Hit rate
Panel B: S&P Ratings																	
74	51	0	0	0	100.00 %	7	7	14	72.55%	1	1	2	96.08%	3	1	4	92.16%
75	73	2	2	4	94.52%	9	6	15	79.45%	1	2	3	95.89%	7	4	11	84.93%
76	88	3	1	4	95.45%	12	12	24	72.73%	2	2	4	95.45%	14	10	24	72.73%
77	96	1	4	5	94.79%	15	7	22	77.08%	1	2	3	96.88%	4	4	8	91.67%
78	219	33	18	51	76.71%	45	32	77	64.84%	29	29	58	73.52%	39	32	71	67.58%
79	221	23	22	45	79.64%	34	32	66	70.14%	30	23	53	76.02%	42	30	72	67.42%
80	209	13	13	26	87.56%	35	24	59	71.77%	30	26	56	73.21%	37	25	62	70.33%
81	249	15	24	39	84.34%	35	26	61	75.50%	24	28	52	79.12%	23	41	64	74.30%
82	191	15	11	26	86.39%	22	19	41	78.53%	18	26	44	76.96%	25	26	51	73.30%
83	229	12	10	22	90.39%	20	15	35	84.72%	14	16	30	86.90%	16	23	39	82.97%
84	233	10	15	25	89.27%	17	25	42	81.97%	21	23	44	81.12%	36	29	65	72.10%
85	265	13	24	37	86.04%	30	37	67	74.72%	46	40	86	67.55%	54	48	102	61.51%
86	293	22	21	43	85.32%	44	30	74	74.74%	35	34	69	76.45%	55	59	114	61.09%
87	304	22	29	51	83.22%	45	35	80	73.68%	55	68	123	59.54%	65	84	149	50.99%
88	271	32	26	58	78.60%	47	41	88	67.53%	36	31	67	75.28%	49	43	92	66.05%
89	204	16	9	25	87.75%	31	20	51	75.00%	21	22	43	78.92%	26	22	48	76.47%
90	145	5	6	11	92.41%	13	11	24	83.45%	12	4	16	88.97%	16	16	32	77.93%
91	155	9	11	20	87.10%	16	19	35	77.42%	17	16	33	78.71%	26	21	47	69.68%
92	183	18	13	31	83.06%	30	21	51	72.13%	24	22	46	74.86%	31	27	58	68.31%
93	198	12	11	23	88.38%	30	20	50	74.75%	27	25	52	73.74%	34	33	67	66.16%
94	194	20	18	38	80.41%	37	28	65	66.49%	27	30	57	70.62%	29	28	57	70.62%
95	211	22	16	38	81.99%	33	32	65	69.19%	33	21	54	74.41%	31	27	58	72.51%
96	213	29	18	47	77.93%	53	36	89	58.22%	43	31	74	65.26%	46	37	83	61.03%
All	4495	347	322	669	85.12%	660	535	1195	73.41%	547	522	1069	76.22%	708	670	1378	69.34%
Avg	195	15	14	29	86.58%	29	23	52	73.77%	24	23	47	78.93%	31	29	60	71.82%

Table 5
Long-Run Performance of Relative Rating Strength Portfolios

On each April 30 from 1974 to 1996, the bond ratings are predicted by four models: multiple discriminant analysis (MDA), multiple discriminant analysis with cross-validation (MDA-C), ordered probit (Probit), and order probit with stepwise variable selection (Probit-S). An under-rated bond portfolio, composed of bonds with predicted ratings higher than agency ratings, and an over-rated bond portfolio, composed of bonds with predicted ratings lower than agency ratings, are formed and held for 5 years. The portfolio returns are rebalanced annually for 5 years and compounded starting from year 2. The abnormal returns are computed by subtracting the compounded returns of the corresponding Lehman Brothers bond index from the portfolio compounded returns. The raw (abnormal) returns in the table are the difference of the compounded (abnormal) returns between under- and over-rated bond portfolios. Panel A reports the results based on Moody's ratings, and panel B lists the results based on S&P ratings. The average annual number of bonds for under- and over-rated bond portfolios are shown in the rows of "U #" and "O #", respectively. "N-W avg" is the number-weighted average return of the relative rating strength portfolio, and "E-W avg" is the equal-weighted average return. "p-value" is the significance level of two-side *t*-test based on equal-weighted average returns. The numbers in the parenthesis are total numbers of misclassified bonds across years used to compute returns in raw and abnormal returns, respectively, for different models.

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel A: Moody's Ratings											
MDA	U #	16	13	11	9	7	16	13	10	8	6
(595,580)	O #	15	14	12	10	7	15	13	11	9	7
	N-W avg	0.92%	2.70%	5.79%	9.98%	15.44%	-0.10%	-0.13%	2.79%	8.45%	13.61%
	E-W avg	0.77%	2.97%	5.95%	9.10%	11.89%	-0.26%	-0.58%	2.32%	6.14%	9.52%
	p-value	0.161	0.002	0.001	0.001	0.003	0.439	0.463	0.177	0.187	0.162
MDA-C	U #	25	20	17	14	11	23	19	15	12	10
(1057,1005)	O #	21	19	16	13	10	21	19	16	13	10
	N-W avg	1.34%	2.86%	6.15%	9.33%	11.19%	0.29%	0.86%	2.06%	3.79%	5.61%
	E-W avg	1.54%	2.80%	5.59%	8.27%	9.90%	0.51%	1.04%	1.97%	3.25%	4.80%
	p-value	0.020	0.000	0.000	0.000	0.000	0.159	0.081	0.038	0.023	0.016
Probit	U #	24	20	16	14	13	23	19	15	13	12
(882,856)	O #	22	19	16	13	10	22	19	16	13	10
	N-W avg	0.79%	2.28%	2.02%	3.06%	10.44%	-0.34%	-0.39%	-2.04%	-2.14%	-0.26%
	E-W avg	0.74%	1.96%	1.76%	1.97%	9.57%	-0.29%	-0.58%	-2.09%	-1.56%	-1.01%
	p-value	0.068	0.009	0.055	0.446	0.000	0.374	0.425	0.086	0.350	0.670
Probit-S	U #	23	19	16	13	12	22	18	15	12	10
(1065,1029)	O #	23	20	17	14	10	23	20	16	13	10
	N-W avg	1.31%	2.55%	3.30%	4.51%	10.11%	0.26%	0.47%	-0.02%	1.13%	2.78%
	E-W avg	1.18%	1.92%	2.74%	3.52%	8.35%	0.54%	0.47%	0.37%	1.79%	3.57%
	p-value	0.052	0.010	0.002	0.002	0.003	0.164	0.534	0.695	0.181	0.073

Table 5 - *continued*

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel B: S&P Ratings											
MDA	U #	17	14	12	9	8	16	13	10	8	7
(631,618)	O #	16	14	12	9	7	16	14	11	9	7
	N-W avg	0.81%	1.88%	4.78%	6.67%	9.92%	0.41%	-0.10%	1.29%	1.95%	3.27%
	E-W avg	0.76%	2.03%	5.24%	6.69%	9.10%	0.41%	-0.18%	1.64%	2.93%	4.62%
	p-value	0.317	0.073	0.001	0.008	0.005	0.468	0.822	0.093	0.161	0.096
MDA-C	U #	27	22	19	15	12	24	20	16	12	10
(1144,1068)	O #	23	20	16	13	10	22	19	16	13	10
	N-W avg	2.20%	4.11%	5.63%	11.78%	17.19%	0.50%	0.84%	1.01%	1.29%	1.45%
	E-W avg	2.11%	3.76%	5.23%	11.41%	14.89%	0.82%	1.08%	1.30%	1.67%	2.29%
	p-value	0.043	0.023	0.016	0.015	0.030	0.066	0.120	0.068	0.105	0.060
Probit	U #	27	23	19	16	14	25	21	17	14	12
(1014,961)	O #	26	23	20	16	13	25	22	19	15	13
	N-W avg	0.74%	1.28%	4.64%	7.55%	14.00%	0.02%	-0.69%	-0.48%	1.54%	3.12%
	E-W avg	0.62%	1.28%	3.52%	6.24%	10.86%	-0.15%	-0.84%	-0.33%	2.12%	2.97%
	p-value	0.231	0.317	0.129	0.047	0.004	0.741	0.403	0.746	0.189	0.180
Probit-S	U #	29	25	22	18	16	28	24	20	16	14
(1323,1271)	O #	28	25	21	17	13	28	24	21	17	13
	N-W avg	2.06%	3.64%	5.76%	7.86%	13.06%	0.27%	0.43%	0.55%	0.98%	1.27%
	E-W avg	1.42%	2.46%	3.58%	6.06%	9.83%	0.41%	0.39%	0.64%	1.71%	2.20%
	p-value	0.076	0.057	0.065	0.005	0.002	0.416	0.648	0.472	0.098	0.163

Table 6
Long-Run Performance of Relative Rating Strength Portfolio: Delete Below-Investment Grade Bonds

This table is similar to Table 5. The only difference is that below-investment grade bonds are deleted when under- and over-rated bond portfolios are formed on each April 30. The raw (abnormal) returns in the table are the difference of the compounded (abnormal) returns between under- and over-rated bond portfolios. Panel A reports the results based on Moody's ratings, and panel B lists the results based on S&P ratings. The average annual number of bonds for under- and over-rated bond portfolios are shown in the rows of "U #" and "O #", respectively. "N-W avg" is the number-weighted average return of the relative rating strength portfolio, and "E-W avg" is the equal-weighted average return. "p-value" is the significance level of two-side *t*-test based on equal-weighted average returns. The numbers in the parenthesis are total numbers of misclassified bonds across years used to compute raw and abnormal returns, for different models.

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel A: Moody's Ratings											
MDA	U #	11	9	7	6	5	11	9	7	6	5
(471,471)	O #	14	13	11	9	7	14	13	11	9	7
	N-W avg	0.63%	1.23%	2.70%	8.74%	13.69%	-0.07%	-0.74%	0.47%	6.35%	11.14%
	E-W avg	0.69%	1.53%	3.27%	7.65%	11.01%	-0.03%	-0.81%	0.81%	4.66%	8.48%
	p-value	0.142	0.087	0.013	0.042	0.059	0.925	0.351	0.574	0.286	0.194
MDA-C	U #	15	13	10	9	8	15	13	10	9	8
(799,799)	O #	19	18	15	12	10	19	18	15	12	10
	N-W avg	0.60%	1.03%	2.52%	5.17%	7.17%	0.04%	-0.18%	-0.12%	1.73%	3.59%
	E-W avg	0.95%	1.50%	2.96%	5.11%	6.95%	0.31%	0.35%	0.64%	1.96%	3.85%
	p-value	0.066	0.009	0.003	0.005	0.001	0.283	0.556	0.372	0.073	0.026
Probit	U #	16	14	11	10	10	16	14	11	10	9
(681,681)	O #	20	17	15	12	9	20	17	15	12	9
	N-W avg	0.11%	1.12%	1.16%	2.92%	6.70%	-0.74%	-0.89%	-2.20%	-2.56%	-0.83%
	E-W avg	0.27%	0.96%	1.32%	2.99%	5.90%	-0.47%	-0.70%	-1.77%	-1.41%	0.28%
	p-value	0.487	0.062	0.130	0.027	0.002	0.232	0.186	0.135	0.407	0.912
Probit-S	U #	16	14	12	10	9	16	14	12	10	9
(848,848)	O #	21	18	16	13	10	21	18	16	13	10
	N-W avg	0.66%	1.50%	1.98%	4.39%	7.98%	-0.24%	-0.09%	-0.63%	-0.22%	1.55%
	E-W avg	0.92%	1.16%	1.73%	3.94%	6.98%	0.21%	0.06%	-0.13%	0.70%	2.74%
	p-value	0.112	0.085	0.104	0.086	0.025	0.557	0.926	0.888	0.676	0.212

Table 6 - *continued*

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel B: S&P Ratings											
MDA	U #	11	9	8	7	6	11	9	8	7	6
(482,480)	O #	15	13	11	8	6	15	13	11	8	7
	N-W avg	0.33%	0.94%	2.14%	4.40%	7.91%	0.09%	-0.17%	-0.04%	0.72%	3.01%
	E-W avg	0.43%	1.02%	2.82%	5.07%	8.77%	0.25%	-0.28%	0.32%	1.59%	4.80%
	p-value	0.434	0.322	0.030	0.127	0.039	0.461	0.700	0.607	0.396	0.079
MDA-C	U #	17	14	12	10	8	17	14	12	10	8
(845,841)	O #	20	18	15	12	9	20	18	14	12	9
	N-W avg	0.21%	0.93%	1.90%	3.31%	4.77%	-0.06%	-0.15%	-0.32%	-0.01%	0.96%
	E-W avg	0.75%	1.55%	2.70%	3.79%	4.91%	0.37%	0.31%	0.34%	0.76%	2.34%
	p-value	0.245	0.084	0.011	0.046	0.030	0.314	0.598	0.587	0.438	0.124
Probit	U #	17	15	13	11	10	17	15	13	11	10
(775,772)	O #	24	21	18	14	12	23	21	18	14	12
	N-W avg	-0.05%	0.65%	1.66%	3.18%	5.85%	-0.30%	-0.58%	-0.36%	0.71%	0.83%
	E-W avg	0.03%	0.61%	1.95%	3.83%	5.90%	-0.22%	-0.52%	0.02%	1.80%	2.39%
	p-value	0.940	0.578	0.088	0.040	0.019	0.621	0.573	0.987	0.210	0.179
Probit-S	U #	20	18	16	13	12	20	18	16	13	12
(1049,1045)	O #	26	23	20	16	13	26	23	20	16	13
	N-W avg	0.02%	0.68%	0.87%	2.08%	5.56%	-0.17%	-0.21%	-0.30%	-0.29%	0.04%
	E-W avg	0.19%	0.58%	1.10%	2.59%	5.60%	0.25%	0.12%	0.40%	1.22%	2.00%
	p-value	0.661	0.507	0.195	0.087	0.012	0.589	0.886	0.635	0.382	0.231

Table 7
Long-Run Performance of Relative Rating Strength Portfolio: Investment Grade Bonds for MDA-C and Probit-S Models

On each April 30 from 1974 to 1996, the ratings for all available investment-grade bonds are predicted by multiple discriminant analysis with cross-validation (MDA-C) and ordered probit with stepwise variable selection (Probit-S). An under-rated bond portfolio, composed of bonds with predicted ratings higher than agency ratings, and an over-rated bond portfolio, composed of bonds with predicted ratings lower than agency ratings, are formed and held for 5 years. The portfolio returns are rebalanced annually for 5 years and compounded starting from year 2. The abnormal returns are also computed by subtracting the compounded returns of the corresponding Lehman Brothers bond index from the bond portfolio compounded returns. The raw (abnormal) returns are the difference of the compounded (abnormal) returns between under- and over-rated bond portfolios. The results based on letter rating and refined rating predictions for MDA model are shown in panel A and B, respectively. The long-run performance of the Probit-S model based on letter and refined rating prediction is reported in panel C and D, respectively. The average annual number of bonds for under- and over-rated bond portfolios are shown in the rows of "U #" and "O #", respectively. "N-W avg" is the number-weighted average return of the relative rating strength portfolio, and "E-W avg" is the equal-weighted average return. "p-value" is the significance level of two-side *t*-test based on equal-weighted average returns. The numbers in the parenthesis are total numbers of misclassified bonds across years used to compute raw and abnormal returns, respectively.

Ratings	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel A: MDA-C (Letter ratings)											
Moody's	U #	15	12	10	9	8	15	12	10	9	8
(685,685)	O #	15	14	11	10	7	15	14	11	10	7
	N-W avg	0.77%	1.42%	3.39%	5.74%	7.52%	0.07%	0.00%	0.28%	1.53%	3.24%
	E-W avg	0.87%	1.71%	3.64%	5.73%	7.61%	0.03%	0.15%	0.55%	1.35%	3.18%
	p-value	0.037	0.001	0.000	0.000	0.000	0.937	0.795	0.480	0.162	0.050
S&P	U #	16	14	12	10	9	16	14	12	10	9
(710,706)	O #	15	13	11	9	7	15	13	11	9	7
	N-W avg	-0.11%	0.32%	1.31%	2.21%	2.38%	-0.45%	-0.43%	-0.53%	-0.65%	-0.85%
	E-W avg	0.07%	0.75%	2.06%	3.01%	2.84%	-0.45%	-0.55%	-0.49%	-0.22%	0.06%
	p-value	0.794	0.305	0.078	0.118	0.200	0.260	0.492	0.597	0.831	0.968
Panel B: MDA-C (Refined ratings)											
Moody's	U #	26	21	18	14	12	26	21	18	14	12
(1144,1144)	O #	24	21	18	14	11	24	21	18	14	11
	N-W avg	0.32%	0.47%	1.05%	1.91%	3.21%	0.06%	-0.30%	-0.49%	-0.69%	0.37%
	E-W avg	0.34%	0.88%	1.77%	2.77%	3.61%	-0.04%	-0.18%	-0.17%	-0.44%	0.41%
	p-value	0.271	0.036	0.068	0.056	0.039	0.881	0.692	0.795	0.581	0.713
S&P	U #	26	22	19	15	13	26	22	19	15	13
(1184,1181)	O #	25	22	19	15	12	25	22	19	15	12
	N-W avg	-0.17%	0.07%	0.84%	0.72%	1.53%	-0.27%	-0.32%	-0.07%	-0.20%	0.31%
	E-W avg	-0.01%	0.45%	1.69%	1.51%	1.81%	-0.29%	-0.27%	0.16%	0.02%	0.34%
	p-value	0.955	0.402	0.061	0.232	0.241	0.355	0.634	0.850	0.986	0.773

Table 7 - *continued*

Ratings	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
Panel C: Probit-S (Letter ratings)											
Moody's	U #	13	11	9	8	7	13	11	9	8	7
(591,591)	O #	13	12	10	8	7	13	12	10	8	7
	N-W avg	0.53%	0.60%	2.07%	4.77%	7.42%	-0.27%	-0.61%	-0.71%	1.08%	2.41%
	E-W avg	0.48%	-0.55%	2.22%	5.83%	7.65%	-0.34%	-1.26%	-0.20%	2.40%	3.67%
	p-value	0.418	0.646	0.045	0.039	0.006	0.454	0.218	0.811	0.252	0.093
S&P	U #	17	15	13	11	10	16	15	13	11	10
(775,771)	O #	17	16	14	12	10	17	16	14	12	10
	N-W avg	0.62%	0.87%	1.62%	4.47%	6.96%	0.10%	-0.47%	-0.76%	0.28%	1.25%
	E-W avg	0.73%	0.95%	2.01%	4.75%	7.04%	0.37%	-0.21%	-0.08%	1.46%	3.33%
	p-value	0.102	0.241	0.037	0.001	0.003	0.472	0.776	0.928	0.236	0.075
Panel D: Probit-S (Refined ratings)											
Moody's	U #	34	29	24	20	17	34	29	24	20	17
(1604,1604)	O #	36	32	26	21	16	36	32	26	21	16
	N-W avg	0.26%	0.44%	0.55%	0.88%	2.90%	0.05%	0.02%	-0.30%	-0.68%	0.31%
	E-W avg	0.49%	0.51%	0.97%	1.50%	3.99%	0.14%	-0.03%	-0.03%	-0.29%	1.09%
	p-value	0.153	0.148	0.202	0.192	0.006	0.595	0.934	0.968	0.749	0.344
S&P	U #	39	35	31	26	21	39	35	31	26	21
(1777,1773)	O #	38	34	29	24	19	38	34	29	24	19
	N-W avg	0.30%	0.72%	1.01%	2.06%	3.50%	0.14%	0.38%	0.28%	0.62%	1.14%
	E-W avg	0.52%	0.90%	1.52%	2.59%	4.24%	0.49%	0.63%	0.69%	1.16%	2.27%
	p-value	0.174	0.089	0.055	0.014	0.008	0.240	0.243	0.390	0.162	0.130

Table 8
Long-Run Performance of Relative Rating Strength Portfolio: Stock Returns and Moody's Ratings

On each April 30 from 1974 to 1995, the Moody's ratings of bonds are predicted by four models: multiple discriminant analysis (MDA), multiple discriminant analysis with cross-validation (MDA-C), ordered probit (Probit), and order probit with stepwise variable selection (Probit-S). The bonds with predicted ratings higher than agency ratings are referred as under-rated bonds, and bonds with predicted ratings lower than agency ratings are called over-rated bonds. Firms with either under- or over-rated bond, but not both, are identified. Only one stock is included for each firm in each year. An under-rated stock portfolio, composed of stocks whose corresponding bonds are under-rated, and an over-rated stock portfolio, composed of stocks whose corresponding bonds are over-rated, are formed and held for 5 years. The portfolio returns are rebalanced annually for 5 years and compounded starting from year 2. The long-run abnormal returns are also computed by subtracting the compounded returns of the corresponding Size/BM control portfolio from the stock portfolio compounded returns. The raw (abnormal) returns are the difference of the compounded (abnormal) returns between under- and over-rated stock portfolio. The average annual number of bonds for under- and over-rated stock portfolios are shown in the rows of "U #" and "O #", respectively. "N-W avg" is the number-weighted average return of the relative rating strength portfolio, and "E-W avg" is the equal-weighted average return. "p-value" is the significance level of two-side *t*-test based on equal-weighted average returns. The numbers in the parenthesis are total numbers of stocks across years used to compute raw and abnormal returns, respectively.

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
MDA	U #	14	13	12	12	12	13	13	12	12	12
(457,429)	O #	10	10	10	10	9	10	10	10	9	9
	N-W avg	-4.46%	-9.43%	-11.94%	-22.37%	-35.08%	-4.60%	-7.87%	-11.69%	-22.81%	-32.24%
	E-W avg	-2.64%	-7.79%	-8.99%	-20.69%	-34.98%	-2.77%	-6.29%	-8.27%	-20.41%	-30.38%
	p-value	0.343	0.081	0.247	0.089	0.049	0.349	0.211	0.365	0.165	0.140
MDA-C	U #	20	19	18	17	16	20	19	18	16	16
(801,754)	O #	14	14	13	13	12	14	14	13	13	12
	N-W avg	-5.55%	-9.26%	-16.10%	-28.85%	-31.09%	-5.47%	-6.66%	-15.48%	-25.87%	-25.06%
	E-W avg	-5.03%	-8.28%	-16.10%	-29.14%	-31.73%	-4.84%	-5.88%	-14.79%	-25.67%	-25.69%
	p-value	0.151	0.111	0.057	0.016	0.020	0.184	0.247	0.093	0.037	0.061
Probit	U #	17	16	15	15	14	16	16	15	14	14
(606,558)	O #	15	14	13	13	12	15	14	13	13	13
	N-W avg	-4.76%	-14.34%	-25.35%	-35.36%	-43.10%	-3.82%	-14.88%	-23.09%	-35.20%	-40.94%
	E-W avg	-5.12%	-14.02%	-21.03%	-33.44%	-45.00%	-3.84%	-13.21%	-18.09%	-30.82%	-41.15%
	p-value	0.115	0.002	0.012	0.001	0.000	0.186	0.002	0.031	0.002	0.003
Probit-S	U #	15	14	14	13	13	15	14	14	13	13
(700,649)	O #	15	14	14	13	13	15	14	14	13	12
	N-W avg	-5.27%	-18.76%	-30.51%	-46.90%	-59.66%	-3.80%	-15.21%	-22.03%	-35.63%	-50.71%
	E-W avg	-6.63%	-19.45%	-29.22%	-46.08%	-60.33%	-4.18%	-13.64%	-18.59%	-32.79%	-50.00%
	p-value	0.025	0.001	0.001	0.000	0.000	0.089	0.012	0.020	0.005	0.001

Table 9
Long-Run Performance of Relative Rating Strength Portfolio: Stock Returns and Discard Firms Below-Investment Grade Bonds

This table is similar to table 10. The only difference is that stocks whose corresponding below-investment grade bonds are deleted when under- and over-rated stock portfolios are formed on each April 30. The raw (abnormal) returns are the difference of the compounded (abnormal) returns between under- and over-rated stock portfolio. The average annual number of bonds for under- and over-rated stock portfolios are shown in the rows of "U #" and "O #", respectively. "N-W avg" is the number-weighted average return of the relative rating strength portfolio, and "E-W avg" is the equal-weighted average return. "p-value" is the significance level of two-side *t*-test based on equal-weighted average returns. The numbers in the parenthesis are total numbers of stocks across years used to compute raw and abnormal returns, respectively.

Model	Statistics	Raw Returns					Abnormal Returns				
		Year 1	Year 2	Year 3	Year 4	Year 5	Year 1	Year 2	Year 3	Year 4	Year 5
MDA	U #	9	9	8	8	8	9	8	8	8	8
(348,328)	O #	10	9	9	9	9	9	9	9	9	9
	N-W avg	-2.03%	-6.30%	-9.84%	-20.93%	-32.88%	-2.39%	-8.07%	-12.48%	-22.70%	-33.44%
	E-W avg	-0.86%	-6.31%	-9.76%	-22.95%	-35.90%	-1.29%	-7.87%	-12.00%	-24.30%	-34.57%
	p-value	0.758	0.138	0.213	0.052	0.023	0.648	0.109	0.179	0.075	0.060
MDA-C	U #	12	11	11	10	10	11	11	11	10	10
(563,533)	O #	13	13	12	12	11	13	12	12	12	11
	N-W avg	-1.86%	-6.32%	-15.07%	-27.04%	-30.18%	-1.90%	-6.51%	-15.13%	-26.83%	-27.53%
	E-W avg	-1.62%	-5.31%	-15.09%	-28.59%	-31.13%	-1.42%	-4.29%	-13.73%	-26.54%	-26.19%
	p-value	0.480	0.240	0.040	0.007	0.005	0.545	0.335	0.063	0.007	0.023
Probit	U #	11	10	10	10	10	11	10	10	10	10
(447,416)	O #	13	12	12	12	11	13	12	12	11	11
	N-W avg	-3.09%	-13.06%	-24.54%	-34.59%	-38.23%	-2.92%	-16.11%	-25.20%	-34.92%	-39.06%
	E-W avg	-3.53%	-13.48%	-21.75%	-31.74%	-38.10%	-3.14%	-15.93%	-22.27%	-31.51%	-37.56%
	p-value	0.165	0.021	0.008	0.004	0.018	0.193	0.010	0.012	0.008	0.038
Probit-S	U #	9	9	9	9	9	9	9	9	9	9
(529,496)	O #	14	13	13	12	12	13	13	13	12	12
	N-W avg	-5.12%	-16.22%	-27.87%	-42.62%	-53.33%	-4.21%	-16.25%	-25.34%	-38.57%	-51.32%
	E-W avg	-6.53%	-18.19%	-29.00%	-42.80%	-52.58%	-4.42%	-14.94%	-22.11%	-33.58%	-44.44%
	p-value	0.040	0.003	0.000	0.000	0.001	0.112	0.008	0.003	0.005	0.005