

BUFN 745 Group Project

Spring 2023

Professor Alex He

Report Due Tuesday, 3/14/2021 at 12pm EST

Computations: Please use Python and Google Colab (or Jupyter) and upload your .ipynb file on Canvas with your submission.

Deliverables: Please submit your slides and a Colab notebook with detailed comments and analysis on Canvas. You will have to prepare a 8 minute presentation of your findings. You should prepare no more than seven slides for your presentation. The presentations will take place on Tuesday, March 14, regular class time. You can find your group number and group members in the spreadsheet “BUFN 745 Final Project Groups” on the shared Google Drive folder.

Setup

You work at the Office of Credit Ratings at SEC and your boss charges you with the task of evaluating the credit rating agencies for corporate bonds. Credit rating agencies face widespread criticism regarding the quality of their ratings. During the financial crisis of 2008-2009, mortgage-backed securities, originally rated AAA, were rapidly downgraded or experienced substantial losses, suggesting that these securities had inflated ratings. For this assignment, you will study the ratings of corporate bonds, which received less attention than mortgage-backed securities but comprises 22% of bonds outstanding in the US.

You have a dataset of S&P Long-Term Issuer Level rating obtained from the ADSPRATE database (“rating.xlsx”). This database contains monthly firm-level rating data starting in 1985. The letter ratings (“*rating*”) are converted into numerical equivalents (“*rt5*”) using an ordinal scale ranging from 1 for the lowest rated firms (CCC) to 17 for the highest rated firms (AAA).¹ The ratings are matched to firm-level accounting data from the latest Compustat entry available before the year-end. (you can find descriptions for each variable in “variable_list.xlsx”)

Questions

Part 1

Your first task is to understand what determines corporate bond ratings. In the spreadsheet “variable_list.xlsx” you can find a list of accounting variables at the firm level that might affect the riskiness of corporate bonds. The goal is to understand what variables can help forecast ratings of corporate bonds.

Please split the sample into train and test parts (do not split by year). Although I have provided you with the data and done some manipulations, there are many ways you can manipulate the data to improve your model. For example, you can winsorize the continuous variables. You might think non-linearities might matter and use various non-linear methods, or include polynomials of X as your regressors for linear methods.

Please be precise in answering the following questions:

¹Because there are very few observations with a CCC rating, we pool them together (CCC–, CCC, CCC+) to form the lowest ordinal category.

1. Explain any data manipulations you may have performed, if any. Which predictors do you use? Did you expand the set of predictors? Explain how and why if you did.
2. Use **at least two** different models to predict ratings. At least one of the models should be multinomial logit, ordered probit, or ordered logit.
3. Which model works best for prediction in sample and out of sample? You can measure error using the mean squared error of the numerical ratings (for example, if the predicted *rt5* is 8, and actual *rt5* is 10, then the error is 2). Which variables are important in determining the bond ratings?

Part 2

Your second task is to measure the changes in rating standards over time. Your boss is considering proposing new regulations for corporate bond ratings, and he needs your input on whether the rating standards have become more or less stringent since 1985. He gives one procedure to test this (You don't have to follow his procedure if you have something better in mind. Be creative here!):

- Estimate a model for ratings using data in 1985;
- Fit the ratings model in 1985 from the previous step to every year after 1985, and calculate the difference between the predicted ratings from 1985's model and the actual ratings. Use the average difference (in numerical rating) as a measure of the difference between the stringency of the ratings in 1985 and that year. For example, if the average rating in 2000 is "A" (*rt5*=12), and the average predicted rating from firm financials is "BBB" (*rt5*=9), then the rating standards are lower in 2000 by $12-9=3$.

You can decide which model to estimate for forecasting bond ratings, but you need to justify your choice. Your boss wants you to be precise in answering the following questions:

1. Did the ratings standards become more stringent from 1985 to 2007?
2. If we separate the sample into speculative grade ("BB+" and below) and investment-grade ("BBB-" and above), are the patterns different from the two types of bonds?
3. If only looking at bonds that are issued for the first time (*ratingage* equals to 0), how does the ratings standards change?
4. Try using a different year other than 1985 to fit the model, and look at the differences between predicted rating and actual rating for other years. Does your conclusions change for whether ratings standards become more stringent from 1985 to 2007?

This list is not exhaustive. Feel free to add items if you find interesting results.

Part 3

Your boss wants to know whether rating standards have real effects on loan outcomes. Your final task is to examine whether rating standards matter for default rates. In other words, for the same firm, if credit rating agencies have looser standards for ratings, is the firm more likely to default?

You have a dataset of average default rates by bond type and by year ("default_by_rating.csv"). Can you conclude whether looser rating standards lead to higher default with this dataset and your estimates from Part 2? Test the hypothesis if you can, or explain why you can not test it.