Faculty of Economic Sciences

Master program M 38.04.08 2022 Economics and Economic Policy

Master program M 38.04.01 2022 Strategic Corporate Finance

Advanced Econometrics

2022-2023

Homework 3 (HW3)

Binary Choice and Ordered Response Models

Instruction

Deadline is 10th of March 2023, 23:59

Final evaluation of Homework 3 is based on **sent materials** and a **presentation** made during seminar.

You should attach materials by Google Form:

Team work

We encourage you to prepare your homework in a team. The goal is to excel your communicative skills and practice in organization of teamwork, to give you opportunity to discuss ideas of econometric analysis and explain econometric routines.

On this task you should work in teams. You can change the list of team participants in the table or keep it the same as in HW 1 and 2.

Decide on the team members and provide the information in the table before the **20th of February.**

We, therefore, expect that each individual contribution is described in short in HW3.

Presentation should be prepared by all team participants. Evaluation of presentation depends on showed knowledge and understanding of each student in the team.

The prepared material should be sent to your tutor via Google Form by one of the team member before the deadline.

HW3 Problem and tasks

Study data and methodology of bank rating. Build binary and ordered response models that allow you to predict the score for a bank based on some bank characteristics. Propose hypotheses concerning the models and test them applying approaches you have learned in the course. Discuss limitations of the study and make conclusions.

Data

Data are given in the folder Data for HW3. Choose the file for your team and download it in *dta* or *xls* format.

In the file you will find date on the ranking positions of some Russian banks and their characteristics.

Methodology

Ratings are opinions about the creditworthiness of a bank. They reflect both quantitative assessments of credit risk and the expert judgment of a ratings committee. Ratings convey information about the relative and absolute creditworthiness of a bank. Agencies regularly publish studies that convey the historical association of ratings and indicators of absolute creditworthiness, such as default ranks and the magnitude of losses at default. Moreover, in the case of structured finance products, ratings are explicitly tied to estimates of default probabilities and credit losses.

Long-term bank ratings are in the 'AAA to D' categories, with '+' and '-' as additional sub-categories. It represents a credit opinion on a bank's ability to meet its contractual financial commitments on a timely basis, and in full, as a going concern. As such, it aims to signal the relative risk of a default-like event. Rating methodology considers that the bank's business model that drives the bank's risk-return strategies, growth and diversification, as well as the way it is seen by its main stakeholders, customers, regulators, creditors, shareholders, counterparties.

To determine the overall qualitative ranking, criteria chosen by the agency are assigned a score from 1 to 100 and the weighted scores are aggregated. Assets, Bank Capital, Credit Portfolio and Net Income are among the key indicators for assigning the score.

Suggested Structure of the Analysis

For each following step briefly describe the procedure and give an interpretation of the results.

I. Data preparation

1. Use the variable **Ranking** to generate a binary variable which is equal to 1, if the bank is ranked with A and higher, and equals 0, if the bank has a lower ranking value.
2. Suggest which explanatory variables influence the rank of a bank. Explain your choice.
3. Investigate the data you are going to use. Provide descriptive statistics with your comments. Account for the outliers if you find some.
4. State hypotheses on influence of assets, bank capital, net income, and their dynamics on the position of a bank in the ranking.

II. Binary choice model

1. Estimate the linear probability model (LPM) by ordinary least squares. Give two reasons why this is not really an appropriate model and what are the advantages LPM gives compared to probit or logit model.
2. How would you interpret positive/negative β coefficients in the model?
3. What are the marginal effects for the linear model?
4. Estimate the model by a **probit** and **logit** model.
5. To what extent is a logit model different from the probit model?

For 7-11 model the probability model as **probit** or **logit** model. Please state which model you choose.

6. Explain the difference between LPM and your (logit or probit) model.
7. Test your model against the model with only intercept. Describe the test you apply.
8. Test the significance of the chosen explanatory variables.

9. Plot the ROC curve. Assess the quality of your model in terms of quality of predicted positive and negative outcomes, specificity and sensitivity.
10. What other criteria can you use to compare the model with the model on constant only? Calculate other measures which allow to assess the quality of a model.
11. Modify the model by removing/adding regressors, including interactions and other nonlinear forms of the explanatory variables. Test and describe the hypotheses related to the steps.
12. Based on the tests conducted above, **propose the final model**. Provide the discussion on the quality of the final model.

For 13-17 choose an explanatory variable which in your opinion is the most important ranking indicator and perform next steps of the analysis.

13. Interpret a positive/negative β coefficient for the chosen variable in the model?
14. Test the hypothesis that the effect of this variable is quadratic.
15. Test whether the effect of this explanatory variable is the same for two groups: banks of Moscow and Moscow region and banks in the rest of Russia.
16. Compute the marginal effect at means. Compute the average marginal effect of the chosen variable on rating position.
17. Based on your analysis make general conclusion on the role of Assets, Bank Capital, Net Income, and their dynamics on the position of a bank in the ranking.

III. Ordered response model

Use the variable **Ranking** to generate an ordered variable. For example, it can be equal to 0 if a bank has the lowest score, 1 – if the bank has a medium score and 2 – if it has the highest score. Describe your choice.

Estimate the probability model as the **ordered probit** or **ordered logit** model. Please state which model you have chosen.

1. Estimate the model.
2. To what extent are ordered logit and probit models different?
3. Test significance of each explanatory variable in your model.
4. Test your model against the model with intercept only.
5. Modify the model by removing/adding regressors, including interactions and other nonlinear forms. Test relevant hypotheses.
6. Based on the tests conducted above, **propose the final model**. Provide the discussion on the quality of the final model.

For 7-9 choose one explanatory variable which in your opinion is the most important ranking indicator and perform next steps of the analysis:

7. Interpret β coefficient for the chosen variable in the model?
8. Compute the marginal effect at means. Compute the average marginal effect of the chosen variable on rating position.
9. Based on your analysis make general conclusion on the role of Assets, Bank Capital, Net Income, and their dynamics on the position of a bank in the ranking.

HW Report Structure

HW Report should contain

1. PDF file with a report.
    a. Report must contain title page with names of all students, group numbers, and the topic of study.
    b. On the first page describe the contribution of each student in the team in detail.
    c. Use Times New Roman-12 font or similar.
    d. The report should contain 10 pages including title page at most. If it contains more pages, the first 10 pages will be reviewed and the paper will be valued with respect to these pages.
    e. All graphs and tables you include in the report should be titled and commented and be useful for understanding data features or the results. If any graphs or tables are useless your score will be decreased.
    f. The results of model estimation should be presented in a table with the following structure. Use "est tab", "outreg2", "put" in STATA or "stargazer" in R. Put statistics for the chosen tests below the estimates of the coefficients. Below the table provide a brief explanation of the tests.

Table: Recommended Structure for Results Presentation

	Name of the model 1	Name of the model 2	Name of the model …
Name of explanatory variable 1	Coef*** (st.error)	…	…
Name of explanatory variable 2	…	…	…
Name of explanatory variable …	…	…	…
Statistics: ● N (# observations) ● Test for joint significance (LR/Wald + p-value) ● McFadden R^2 ● LogLikelihood ● AIC, BIC ● …			

*, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively

2. Attach to the report Do-file / R script or log-file (Stata / R file with commands or commands and output).

Presentation

Presentation itself is expected to last for 5-7 minutes. It should contain basic approaches, hypotheses and the results of your study. Your results will be discussed together with students and the tutor for 5-7 minutes after presentation. You are expected to answer questions related to your presentation. Activity of participation in discussion is evaluated.

Dates for presentation will be announced by the seminar instructor.

You can choose the date of presentation which is convenient for your team. Be aware that the number of teams that can choose the same seminar for presentation is limited by duration of the seminar and other tasks to be solved during seminar.

If a participant is absent on the date of presentation, he or she has to come for presentation during tutor's office hours. HW3 will be evaluated on the base of the main and additional presentations of all participants in the team.

Final Evaluation of HW3 (FE)

$$FE=0.5HW\_Report + 0.5HW\_Presentation$$

HW_Presentation = average (Individual Scores for presentation made by team participants)

HW_Report is evaluated on the base of the structure given above.

If you have any questions do not hesitate to contact your tutor.

Good luck!