MS&E 246: Financial Risk Analytics

Course project: Design, implement and test a model of small-business loan charge-off ("default") based on a data set of roughly 150,000 equipment loans backed by the US Small Business Administration (SBA) between 1990 and 2014, which will be provided as an Excel spreadsheet. Make sure you understand the nature of the loans (see the SBA website for more details on the loan program, which is called "504"), the structure of the data set and the variables provided.

- Explore the loan data set to inform model building.
- Think about the treatment of missing data.
- Justify your model and the choice of predictor variables. Explore predictor variables beyond those provided in the SBA data set, in particular time-varying risk factors at the regional (i.e. zip-code) level. Also consider linear and nonlinear model alternatives.
- Select and implement an appropriate method for fitting the model parameters.
- Rigorously test predictive performance (in- and out-of-sample) of your model alternatives using a Receiver-Operating-Characteristic (ROC) curve and other appropriate metrics.
- Explain the fitting results and the fitted model. Which variables are important and why?
- Next develop, fit, and evaluate a model for the loss at default.
- Then, estimate the distribution of total loss on a portfolio of 500 randomly selected loans over one and five year periods (state the loan selection method). Measure the risk in terms of the VaR and the Average VaR (also known as expected shortfall) at the 95% and 99% levels (include confidence bands for your estimates).
- Finally, estimate the distributions for the one and five year losses of an investor who has bought a [5%,15%] tranche backed by the chosen portfolio. Also consider a [15%, 100%] senior tranche. Interpret and compare the distributions from a risk management perspective.

Write up a final report, detailing your models, statistical estimation approaches, tests, and results to the extent that the results could be replicated.

Please form a team of up to 4 students by Tuesday 1/11/22 and send an email with the team composition (including names and email addresses) to the TA (one email per team).

The teams will present their results in the last class on March 10, 2022.