MS-E2177 Seminar on Case Studies in Operations Research

S-Bank: Allocation of the Sales Price of the Credit Collection Portfolio

Project plan

Joonatan Honkamaa, Project Manager
Juhana Vehmas
Mattias Österbacka
Petri Koivisto

March 27, 2023

1 Background

S-bank has 3.2 million customers and according to preliminary figures an operating profit of almost 45 million EUR in 2022 [1]. A significant part of S-bank's business is related to lending activities, having a comprehensive collection of different loans and credits. Examples of these are secured loans such as mortgages, and unsecured loans such as "S-laina" and card credits.

The life cycle of each credit starts from identifying and pre-qualifying new customers. For these potential new customers loan offers are created. If an agreement is made, the loan enters the Account Management phase. If everything goes according to plans and the repayment plan is fulfilled by the customer, the life cycle of a loan comes to an end. However, in the Account Management phase, there might occur some problems with the payments, and the loan enters the Pre-delinquency Collections phase where the loans of customers, which have the probability of default in the near future, are proactively managed. In the occurrence of late payments, the bank starts collections without legal action (soft collection).

If the payments are late for over 90 days, a default occurs. After the event of default, the bank can still try to collect recoveries by themselves and/or sell the defaulted loan to a collection agency.

This project focuses on the phases after the event of default, and more specifically the scope is in the loans that have been sent to a collection agency. The bank can assess the risks of giving loans by calculating expected loss (EL). Expected loss can be calculated with

$$\mathrm{EL} = \mathrm{PD} \cdot \mathrm{EAD} \cdot \mathrm{LGD}, \tag{1}$$

where

- PD is the probability of customer's default,
- EAD is exposure at default, i.e. the balance of the loan in EUR at default,
- LGD is loss given default, i.e. percentage of the loan that bank is not able to collect after default.

Banks are allowed to calculate their own risk parameters, this is known as internal rating-based approach (IRBA). There are two types of IRBAs that can be applied, foundation IRBA, and advanced IRBA. Both allow the banks to calculate their own probability of default (PD). The advanced IRBA also allows banks to estimate exposure at default (EAD), and loss given default (LGD).[2] This process requires supervisory regulation and approval [3].

The main objective of this project is to estimate LGD, i.e. the financial loss a bank ultimately incurs when a borrower stops making loan payments. The LGD value is expressed as a percentage of the bank's total exposure at the time when a borrower defaults. LGD is in practice estimated based on previous observations of LGD-values of loans. However, data is not easy to collect because the actual costs of a defaulted loan are often scattered over many parts, some of which are not easy to track. On top of that, the bank does not receive any information on loans sold to collection agencies, and thus the estimation has to be made based on total price of the portfolio.

Without predictive information LGD can be estimated by just taking mean of the observed values. In practice, however, the bank has a lot of information about every loan which can be used to predict the LGD-value. A simple approach to predict the value is to divide the loans to different groups, and use the mean of the group as an estimate. Another way is to construct regression models. Different types of regression models include linear regression, Tobit regression, beta regression [4], inflated beta regression and censored gamma regression. [5] The two approaches mentioned above can also be combined.

Third approach is to use a two-stage model, where the probability of LGD-value being zero is first estimated, and a regression model is applied to data which only has LGD-values larger than zero. [6]

2 Objectives

The aim of the project is to analyze the recovery cash flows from different points in time and to estimate potential cash flows for collection portfolios that have been sold at different points of collection.

The case team is expected, following the guidance and data provided from S-Bank,

1. to analyze different allocation methods for the portfolio sale price to individual loans for LGD-modelling purposes.

In order to do that the team should

2. develop a justified and documented model to estimate a simulated cash flow for loans that have been sent to collection agency,
3. investigate the effect of different characteristics of the loan and lender to the sale price of the collection portfolio,
4. analyze potential effects of asymmetrical information to the sale price of collection portfolios (collection agencies have better picture of the lenders overall financial situation).

3 Tasks

We have a large dataset of loans that have defaulted. We will familiarise ourselves with this data. A model that predicts future cash flows for defaulted loans will be developed. We use statistical methods to find explanatory variables for the future payments. The portfolio sale price will then be allocated for the individual loans based on the simulated cash flows using an optimization model. Other methods for allocating will also be explored.

4 Schedule

The tentative schedule for the project is presented below in Figure 1.

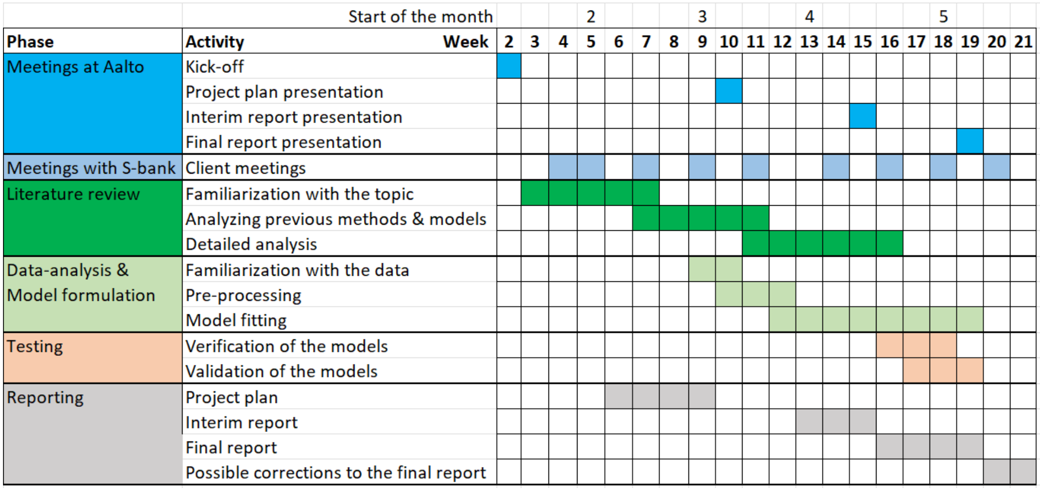

Figure 1: Gantt chart for the project.

5 Resources

Our project team consists of four Systems and Operations Research students. This means that all members have skills in programming, mathematics and statistics. The downside, however, of the group is, that nobody has significant previous experience in finance, meaning that we need to do our background work properly before moving on to the actual modelling. Our project manager, Joonatan Honkamaa, will make sure that everyone does their background work properly and that the tasks of the project will be distributed evenly.

Our main contact at S-bank is Petri Viertiö provides us with analytical assistance and general guidance throughout the project. Petri is chief risk officer at S-bank and has a large amount of experience from optimisation models related to finance and banking. The course teacher Professor Ahti Salo is a general supervisor and offers guidance to issues related to the course.

Other contacts at S-bank are Karri Holopainen, who is the main credit risk manager at S-bank, and Elina Tuomi, data analyst, who collects the data for the project. The data in use consists of about 30 000 minor loans of two types, “Type 1” and “Type 2”, and their information. The data of a single loan starts at the moment when the loan is transferred to collection, and ends to the moment when no more information is available, for example when the loan is sold to collection agency.

6 Risks

Risk	Probability	Effect	Impact	Prevention
Poor data quality	High	Created models are not predictive	High	Active communication with S-bank experts, identification of outliers and justified assumptions
Team member inactivity	Low	Increased workload of other members	Medium	Clear allocation of tasks and investing to building team spirit
Model overfitting	Medium	The model reacts too strongly to data features	Medium	Careful analysis of model performance
Macroeconomic phenomena have affected the data	Medium	The model does not represent current behavior of loan cash flows	Medium	Recognition and clear communication about the issue
Insufficient communication with the client	Medium	The model does not satisfy the requirements	High	Clear and regular communication between the team members and the client
Predictive power of the model is poor	Medium	The results are not useful	Medium	Analysis of model ideas within our team and with the client

Table 1: Recognized risks of the project.

References

[1] S-pankki.fi website Retrieved on 14.02.2023. https://www.s-pankki.fi/fi/tiedotteet/2023/positive-profit-warning-s-bank-groups-operating-profit-in-2022-exceeds-forecast

[2] Risk.net website. Retrieved on 13.02.2023. https://www.risk.net/definition/internal-ratings-based-irb-approach

[3] Bundesbank.de website. Retrieved on 13.02.2023. https://www.bundesbank.de/en/tasks/banking-supervision/individual-aspects/own-funds-requirements/credit-risk/internal-ratings-based-approach-622830

[4] Huang, Xinzheng & Oosterlee, Cornelis. (2011). Generalized Beta Regression Models for Random Loss-Given-Default. Reports of the Department of Applied Mathematical Analysis, 08-10. 7. 10.21314/JCR.2011.150.

[5] Yashkir, Olga and Yashkir, Yuriy (2013): Loss Given Default Modelling: Comparative Analysis. Published in: Journal of Risk Model Validation , Vol. 7, No. 1 (27 March 2013)

[6] mathworks.com website. Retrieved on 13.02.2023. https://se.mathworks.com/help/risk/comparing-lgd-models.html