PROJECT PROGRESS REPORT
Prepayment Modeling for Small Business Loan Asset-Backed Securities

July 26, 2023

Samuel Kang
400177775

Contents

0.1 Abstract

The issue of prepayment for MBS (Mortgage-Backed Securities) or other forms of Asset-Backed Securities poses a challenge in terms of modeling. It requires a rigorous approach that involves complex interest rate modeling and consideration of various factors. In this project, we propose an alternative analytical approach to address this challenge, which can be divided into two key parts.

Firstly, we aim to reduce the number of existing MBS tickers through clustering techniques. By applying clustering analysis, we can group similar MBS together, leading to a smaller number of larger groups. This approach allows us to analyze the prepayment speed for these groups, providing valuable insights into their consistent behaviors. We anticipate that this clustering analysis will help distinguish potentially four groups exhibiting differing prepayment characteristics.

Secondly, we aim to identify statistically significant attributes that can predict prepayment speed. By exploring various factors, including macroeconomic indicators such as interest rates and inflation, we can determine which attributes have the most significant impact on prepayment. Based on this analysis, we will model these attributes as functions to predict prepayment for each group identified in the clustering analysis.

By implementing this alternative approach to prepayment modeling, our objective is to streamline the modeling process while still capturing the essential factors that influence prepayment behavior. This will enhance our understanding of prepayment dynamics for small business loan ABS and improve our predictive capabilities.

Through this project, we seek to contribute to the advancement of prepayment modeling techniques for small business loan ABS, providing valuable insights to investors and industry professionals. Our goal is to enhance the understanding of prepayment dynamics and improve the accuracy of prepayment predictions for these types of asset-backed securities.

0.2 Project Progress

1. June 9

    - Read through the standard prepayment modeling method, and have a high-level idea.

2. June 12

    - When people aggregated these loans, to generate polls of loans. How are pools generated? Not sure they aggregated by asset type? (ie. asset base on - equipment loan? real estate loan? car loan or mix it all together) Investigated through the SBA website, to find out how the pools were generated.

    - The SBA website indicates The loan pool together by 1. the floating rate they are referring 2. the level of the spread that is below/above the floating rate for example, prime rate + 30/60 bp. One of the characteristics for pooling is the range of the spread (prime + spread in between 30-60 bp), then we can calculate the weighted average coupon.
      Base Rate and Adjustment: For variable rate pools, each underlying Loan must use the same base rate, either the Wall Street Journal Prime Rate, the LIBOR Base Rate2, or the SBA Optional Peg Rate, and float on the same accrual basis, either monthly or calendar quarterly

    - Need to be done: to find if there is a clear indication of the type of underlying loans. investigated from the loan assemblers' side (they will take on SBA loans and make them into pools and securitize them and sell to investors), for example, BMO, FHN

3. June 20

    - Go through sample83164lZJ3.xlsx, understand the CPR (Conditional Prepayment Rate)
    - Build model for 1. define the prepayment speed 2. rank the speed 3. build initial logistics regression for fast/slow(1,0) speed
    - Identifying key factors that impact prepayment speed is crucial as they can have a significant influence on the outcome, based be default modeling prepayment literature.

4. June 29

   - Go through Bloomberg data from the sample ticker screenshot and discuss which variables are useful for our model intuitively.
     - Liquidity Score (BVAC score)
     - Yield
     - Life of the loan (Vintage)
     - Maturity data
     - Original loan amount
     - Current loan amount
     - Original numbers of the loan
     - Current numbers of the loan

5. July 6

   - Once get the data pulled from Bloomberg, general steps for the regression to model the prepayment part:
     (a) Variable Selection. Selection from both statistical analysis and intuition/economic considerations. We will begin by the raw statistics to find variables that show statistical significance. Once we have these variables, we will explore their economic implications and assess whether they align with our intuition through economic research.
     (b) Model Specification. Linear regression with multiple variables? non-linear term? the auto regressive term? error-correcting term? (Analyze from statistical analysis and intuition/economic considerations)
     (c) Testing to make sure the model is sound. We could use FHN/ Bloomberg projection as the benchmark to assess our model's accuracy/sound.
   - Need to be done: 1. seeking some more variables to add to our model whether for the classification part or regression part. For example GDP, HPI, CPI

6. July 12

   - Go through the dataset provided
   - Need to be done:
     (a) We only have one single value of CPR, but we need a time series of CPR. Need more data, full time series of CPR to start regression analysis.

(b) fixes the granularity of the macroeconomic data, taking the average from daily data to quarterly data.

(c) Start doing analysis on the variables relative to each other, for example, for Linear Regression multicollinearity test.

7. July 20

- updated with continuous time-series 6-month CPR data
- Tasks:

(a) covert CPR continuous data to a more representative average data per ticker. Using the normal average, or weighted average method

(b) Look at the distribution for the CPR data and try to find a good threshold for the groups(high/low speed)

(c) Bring the average of the time series into our old data and run the logistics test

(d) Maybe run multiply logistic model by term(short or long) or other type of characteristics

0.3 Model Building

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix,
    accuracy_score

SBA_data = pd.read_excel("SBA_data.xlsx")

x=SBA_data.drop(["CPR_indicator","CPR"],axis=1)
y=SBA_data["CPR_indicator"]

x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3,
    random_state=1)
logmodel=LogisticRegression()
logmodel.fit(x_train,y_train)
predictions= logmodel.predict(x_test)
classification_report(y_test,predictions)
confusion_matrix(y_test,predictions)
accuracy_score(y_test,predictions)
```

Listing 1: Python example

0.4 Results

Main Results related to the purpose of the research or things to disccuss [1] [2] [3] [4] [5] [6]

Bibliography

[1] Christian Ahlin and Robert M Townsend. Using repayment data to test across models of joint liability lending. *The Economic Journal*, 117(517):F11–F51, 2007.

[2] Pan Kang and Stavros Zenios. Complete prepayment models for mortgage-backed securities. *Management Science*, 38:1665–1685, 11 1992.

[3] Eduardo S Schwartz and Walter N Torous. Prepayment and the valuation of mortgage-backed securities. *The Journal of Finance*, 44(2):375–392, 1989.

[4] Richard Stanton. Rational prepayment and the valuation of mortgage-backed securities. *The Review of financial studies*, 8(3):677–708, 1995.

[5] Kenneth Temkin and Roger C Kormendi. An exploration of a secondary market for small business loans. *Washington, DC: Small Business Administration(April). http://www. sba. gov/advo/research/rs227_ tot. pdf*, 2003.

[6] Erin K Walsh. Small business loan abs. *Structured Products and Related Credit Derivatives: A Comprehensive Guide for Investors*, pages 147–156, 2012.