Technocolabs Softwares Inc.

MORTAGE TRADING & MORTAGE LOANS

DONE BY : TEAM A
SUPERVISED BY : Mr. SHAH Yasin
2023/2024

Table of Content

1. Understand Prepayment Dynamics

i) Research and define mortgage trading.

The factors influencing prepayment behavior in mortgage loans:

1. Interest Rates

- **Interest Rate Levels:** One of the primary drivers of prepayment is the level of interest rates. When interest rates drop, borrowers are more likely to refinance their existing mortgages to take advantage of lower rates, leading to an increase in prepayments. Conversely, higher interest rates tend to decrease prepayment activity.
- **Interest Rate Spread:** The difference between a borrower's current mortgage rate and the prevailing market rates can influence the decision to prepay. A wider spread creates a stronger incentive for refinancing.

2. Borrower Characteristics

- **Credit Score:** Borrowers with higher credit scores are generally more likely to prepay their loans, as they are more likely to qualify for refinancing at better rates.
- **Income Level:** Higher-income borrowers may prepay their loans more frequently, either through refinancing or by making larger-than-required payments, due to better financial flexibility.
- **Employment Status:** Stable employment increases the likelihood of prepayment, while unemployment or underemployment can reduce a borrower's ability to prepay.
- **Age:** Younger borrowers may prepay less frequently, as they are often in the earlier stages of their mortgage term and have less equity. Older borrowers, especially those nearing retirement, may be more inclined to prepay to eliminate debt.

3. Loan Terms

- **Fixed-Rate vs. Adjustable-Rate Mortgages (ARMs):** Borrowers with ARMs may be more prone to prepayment, particularly if their adjustable rates are set to increase. In contrast, those with fixed-rate mortgages may be less motivated to prepay if their rates are lower than the current market rates.
- **Loan Tenure:** The stage of the loan affects prepayment behavior. Borrowers are more likely to prepay during the early years of a mortgage when they may be refinancing or moving to a new home. As the loan matures, prepayments tend to decrease.
- **Loan-to-Value (LTV) Ratio:** A lower LTV ratio generally correlates with higher prepayment rates, as borrowers with more equity in their homes are better positioned to refinance or pay off their loans.

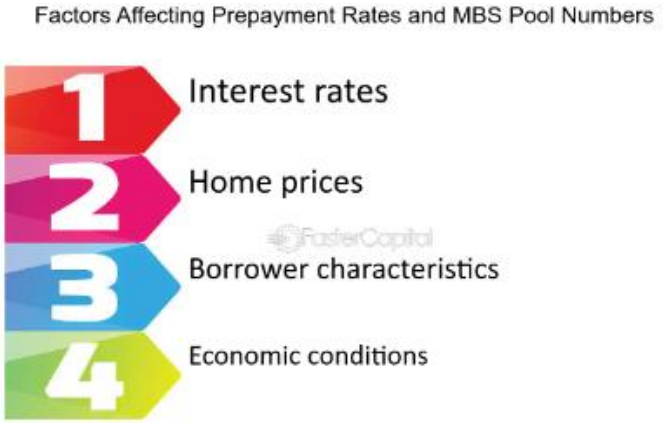
4. Macroeconomic Conditions

- **Economic Growth:** During periods of economic expansion, increased employment and rising incomes can lead to higher prepayment rates. Conversely, economic downturns often reduce prepayment activity due to financial uncertainty.
- **Inflation:** Higher inflation can lead to higher interest rates, reducing the incentive for borrowers to refinance or prepay their loans.
- **Housing Market Conditions:** A strong housing market with rising home values can lead to increased prepayments as borrowers refinance or sell their homes. In a declining market, prepayments may decrease as borrowers are less likely to move or refinance.

- **Government Policies:** Regulations and government programs, such as tax incentives for refinancing or modifications to mortgage interest deduction rules, can significantly influence prepayment behavior.

5. Other Factors

- **Psychological and Behavioral Factors:** Borrowers' personal preferences, risk tolerance, and financial planning strategies also play a role in prepayment behavior. Some may prioritize paying off debt quickly, while others may prefer to invest excess cash elsewhere.



ii) Analyze historical prepayment data to identify any trends or patterns.

To understand prepayment dynamics and assess prepayment risk, analyzing historical mortgage loan data is crucial. Here, we analyze a sample of historical mortgage loan data with features such as credit score, first payment date, occupancy, and loan performance metrics. This analysis aims to identify trends and patterns in prepayment behavior based on the provided dataset.

a) Dataset Overview and Observations:

The dataset provides various features related to mortgage loans, including borrower credit scores, mortgage insurance premiums (MIP), loan-to-value ratios (LTV), debt-to-income ratios (DTI), and loan performance metrics. Below is a detailed analysis based on the provided dataset and additional remarks.

- **Factors Influencing Prepayment Behavior**

Distribution: The majority of borrowers have credit scores ranging between 500 and 800. A significant number of borrowers, however, have scores ranging from 0 to 19. This unusual range may indicate data anomalies or errors, suggesting a need for further investigation or data cleaning.

Implications: High credit scores generally correlate with a higher likelihood of refinancing and prepayment. The presence of anomalous data points needs to be addressed to ensure accurate risk modeling.

- **Mortgage Insurance Premium (MIP)**

Categorical Variable: MIP values predominantly range between 0 and 1, indicating that most loans either have no insurance premium or a basic level of coverage. MIP should be treated as a categorical variable reflecting the type of insurance policy.

Implications: Understanding the MIP categories can provide insights into the risk profiles associated with different insurance policies.

- **Original Combined Loan-to-Value Ratio (OCLTV)**

Distribution: OCLTV values primarily fall between 80% and 84%, with a notable percentage between 95% and 99%. A small percentage of loans have OCLTVs between 20% and 40%.

Implications: High OCLTV ratios indicate a higher risk of lending, as borrowers have less equity in their properties. This suggests a potential higher prepayment risk for high-OCLTV loans

- **Debt-to-Income Ratio (DTI)**

Distribution: A significant number of borrowers have a DTI of 0, suggesting strong financial health. Conversely, many have DTI values around 29% or 30%, indicating a stable income relative to debt.

Implications: Low DTI ratios are associated with lower financial risk, which could influence prepayment behavior as financially stable borrowers might refinance to benefit from lower rates.

- **Property Values**

Distribution: Most properties are valued between $8,000 and $250,000, with a small percentage falling outside this range.

Implications: Property value distribution affects the overall loan size and risk. Loans for high-value properties may have different prepayment dynamics compared to those for lower-value properties.

- **Interest Rates**

Distribution: Interest rates are primarily between 6% and 7.5%, with a few extreme values above 7% or below 6%.

Implications: Interest rates within this range are typical for the dataset. Extreme values may indicate special cases or loans with unique terms.

- **Loan Terms**

Repayment Periods: All loans have a 360-month term, but the repayment periods observed range from 30 to 200 months.

Implications: The concentration of repayment periods between 30 and 75 months suggests commonality in the dataset. Longer periods may represent long-term loans, and shorter periods could indicate early prepayments.

b) Results and Insights

The analysis reveals that most borrowers are not first-time buyers, suggesting that their prepayment behavior may be influenced by previous mortgage experiences. This background could impact their decisions regarding prepayment and refinancing. Additionally, the data shows that the majority of loans are classified as "O" for owner-occupied properties. This classification is important for assessing prepayment risk, as owner-occupied properties often exhibit different refinancing patterns compared to rental properties. Generally, the dataset shows that California is the most frequent property state, indicating a concentration of loans in this region.

This concentration means that regional economic conditions and housing market trends in California will significantly influence prepayment behavior and associated risks.

Most loans have repayment periods between 30 and 75 months, with a median around 50 months, reflecting a typical loan duration. However, some loans extend beyond 200 months, representing long-term loans that may exhibit different prepayment dynamics. Interest rates are predominantly between 6% and 7.5%, with rates outside this range potentially indicating unique loan conditions or anomalies.

This common range aligns with typical market conditions, though variations in interest rates can impact prepayment behavior. Additionally, the distributions of OCLTV (Original Combined Loan-to-Value) and LTV (Loan-to-Value) ratios are similar, suggesting that borrowers rarely take out second mortgages. This similarity indicates a consistent level of equity in the properties.

2. Define Prepayment Target Variable

In developing models to predict prepayment behavior, it's crucial to define the target variable accurately. This variable can either be binary or continuous, depending on the specific objectives of the analysis.

i) Binary Target Variable

- **Definition:** The target variable indicates whether a loan has been prepaid within a specific timeframe. It takes on a value of 1 if the loan was prepaid and 0 if it was not.
- **Timeframe Considerations:**
  - **12-Month Prepayment Indicator:** Prepaid_12M = 1 if the loan is prepaid within 12 months, otherwise 0.
  - **24-Month Prepayment Indicator:** Prepaid_24M = 1 if the loan is prepaid within 24 months, otherwise 0.
- **Use Cases:** This approach is useful when the primary goal is to identify loans that are at risk of prepayment within a given period.

ii) Continuous Target Variable

- **Definition:** The target variable represents the percentage of the loan amount that has been prepaid within a specific timeframe. It ranges from 0 (no prepayment) to 100 (full prepayment).
- **Timeframe Considerations:**
  - **12-Month Prepayment Percentage:** Prepaid_Pct_12M = (Amount Prepaid within 12 Months / Original Loan Amount) * 100.
  - **24-Month Prepayment Percentage:** Prepaid_Pct_24M = (Amount Prepaid within 24 Months / Original Loan Amount) * 100.
- **Use Cases:** This approach is beneficial when it’s important to predict the degree of prepayment, such as for assessing the impact on cash flows or MBS pricing.

iii) Timeframe Considerations

- **Short-Term (e.g., 12 Months):**
  - **Focus:** Captures more immediate prepayment behavior, potentially more responsive to short-term economic changes (e.g., interest rate fluctuations).
  - **Use Case:** Useful for short-term risk assessment and monitoring.
- **Medium-Term (e.g., 24 Months):**
  - **Focus:** Provides a broader view, capturing longer-term prepayment behavior influenced by factors such as borrower refinancing decisions and economic cycles.
  - **Use Case:** Suitable for medium-term forecasting and strategic planning.

1. Calculate the Scheduled Balance:

• This is the balance that should have remained according to the original payment schedule without any extra prepayments.

• The formula to calculate the scheduled balance after a certain timeframe (Months Passed) is:

$$\text{Scheduled Balance} = \text{Orig UPB} \times \left(1 - \frac{(\text{Months Passed} \times \text{Monthly Payment})}{\text{Orig UPB}}\right)$$

Where:

• **OrigUPB** is the original loan balance.

• **Monthly Payment** is calculated based on the monthly interest rate and loan term.

• **Monthly Interest Rate** = Annual Interest Rate / 12.

• **Loan Term** is in months (e.g., 360 months for a 30-year mortgage).

2. Determine the Actual Remaining Balance

• Compare the expected balance with the actual remaining balance after the selected timeframe. If the borrower has prepaid more than expected, the remaining balance will be lower than the scheduled balance.

$$\text{Prepaid Amount} = \text{Scheduled Balance} - \text{Remaining Balance at 12 Months}$$

3. Calculate the Prepayment Percentage

• The formula is:

$$\text{Prepaid Percentage} = \left(\frac{\text{Prepaid Amount}}{\text{Orig UPB}}\right) \times 100$$

• This represents how much more than the scheduled amount has been paid.

Implementation of above concept using sample data:

Orig_UPB	Scheduled_Balance_at_12_Months	Remaining_Balance_at_12_Months	Prepaid_Amount
117,000	110,000	108,000	2,000
109,000	106,000	105,000	1,000
88,000	87,000	86,000	1,000
160,000	152,000	150,000	2,000

Prepaid_Amount	Prepaid_Percentage (Target Variable)
2,000	1.71%
1,000	0.92%
1,000	1.14%
2,000	1.25%

I have used 12 months as a constant time interval we can use any particular time interval as per our requirement.

3. Data Exploration and Feature Engineering

i. Examine the available data to identify relevant features that correlate with prepayment behavior.

- **CreditScore**: Borrowers with higher credit scores might have more refinancing opportunities, potentially leading to higher prepayment rates.

- **FirstPaymentDate and MaturityDate**: The time remaining until maturity can affect prepayment. Loans with a longer time remaining may have a different prepayment profile compared to those nearing maturity.

- **FirstTimeHomebuyer**: First-time homebuyers might behave differently from repeat buyers in terms of prepayment due to different financial situations and incentives.

- **MIP (Mortgage Insurance Premium)**: The presence of mortgage insurance might influence prepayment, as borrowers might refinance to remove MIP once they have enough equity.

- **Units**: The number of units in the property (e.g., single-family, multi-family) could impact prepayment, as different property types might have different risks and incentives.

- **Occupancy**: Owner-occupied properties might have different prepayment behaviors compared to investment properties.

- **OCLTV (Original Combined Loan-to-Value)** and **LTV (Loan-to-Value)**: These ratios indicate how much equity a borrower has in the property, which can be a significant factor in prepayment decisions.

- **DTI (Debt-to-Income Ratio)**: Borrowers with higher DTI ratios may be more sensitive to interest rate changes, influencing their likelihood to refinance and prepay.

- **OrigUPB (Original Unpaid Principal Balance)**: The loan amount could correlate with prepayment, as borrowers with higher balances might be more likely to refinance to save on interest payments.

- **OrigInterestRate**: The original interest rate is a critical factor in prepayment, as borrowers are likely to refinance when market rates drop below their original loan rate.

- **Channel**: The loan origination channel (e.g., retail, broker, correspondent) might affect prepayment behavior due to different borrower profiles and loan terms.

- **PPM (Prepayment Penalty)**: The presence of a prepayment penalty can deter borrowers from prepaying early, thus impacting prepayment rates.

- **ProductType**: The type of mortgage product (e.g., fixed-rate, adjustable-rate) can influence prepayment behavior, with adjustable-rate mortgages (ARMs) potentially leading to higher prepayment rates as borrowers refinance when rates adjust.

- **PropertyState** and **PropertyType**: Geographic location and property type can impact prepayment, with certain states or property types (e.g., condos vs. single-family homes) having different prepayment risks.

- **LoanSeqN (Loan Sequence Number)**: This might help track individual loan performance over time.

ii. Examine the Create derived variables (e.g., loan-to-value ratio, interest rate spread) that may act as predictors.

Loan-to-Value (LTV) Ratio: Measures the loan amount relative to the property's value.

$$\text{LTV} = (\text{Loan Amount} / \text{PropertyValue}) * 100$$

- **Loan Amount**: The initial principal balance of the loan.
- **PropertyValue**: The appraised value of the property at the time the loan was originated.

Interest Rate Difference: To capture the difference between the current and original interest rates:

$$\text{Interest Rate Difference} = \text{Current Interest Rate} - \text{Original Interest Rate}$$

- **Current Interest Rate**: The prevailing market interest rate at the time of analysis.
- **Original Interest Rate**: The interest rate at the time the loan was originated.

- **Time Since First Payment**: To capture the time elapsed since the first payment was made:

$$\text{Time Since First Payment (in months)} = \text{Current Date} - \text{First Payment Date}$$

This can be calculated in months or years,

Loan-to-Value (LTV) Ratio: the LTV ratio is already provided as part of the dataset (the LTV feature).

Interest Rate Spread: This measures the difference between the interest rate on the mortgage and the prevailing market interest rate :

$$\text{Interest Rate Spread} = \text{Original Interest Rate} - \text{Market Interest Rate}$$

- **Original Interest Rate**: The interest rate on the mortgage at the time of origination (from your dataset).
- **Market Interest Rate**: The prevailing interest rate for similar loans at the time of origination or at the time of analysis.

- **Equity Buildup**:

$$\text{Equity Buildup} = \text{Current Property Value} - \text{Remaining Loan Balance}$$

Captures the borrower's equity in the property, which may influence their decision to prepay.

4. Against Industry Standards

i) Review existing models and research papers on prepayment modeling in the MBS sector to identify commonly used methodologies.

a. Hazard Models (Survival Analysis)

- **Overview**: Hazard models, such as Cox Proportional Hazards, are frequently used to estimate the likelihood of prepayment over time, taking into account various covariates like interest rates and borrower characteristics.
- **Key Research**: Search for studies that apply survival analysis to mortgage datasets, focusing on how these models capture time-dependent risk factors influencing prepayment.

Survival models relate the time that passes, before some event occurs, to one or more covariates that may be associated with that quantity of time.

b. Logistic Regression

- **Overview**: Logistic regression models are used to predict binary outcomes, such as whether a loan will be prepaid or not. This approach is relatively simple and interpretable, making it a common choice in prepayment modeling.
- **Key Research**: Look for papers that use logistic regression to model the impact of features like loan-to-value ratios, credit scores, and economic indicators on prepayment behavior.

c. Machine Learning Models

- **Overview**: Advanced machine learning techniques, including Random Forests, Gradient Boosting Machines, and Neural Networks, are increasingly used for prepayment modeling. These models can capture complex nonlinear relationships between features.
- **Key Research**: Identify research that applies machine learning models to prepayment data, emphasizing their predictive performance and feature importance analysis.

d. Option-Theoretic Models

- **Overview**: These models treat prepayment as a financial option that borrowers exercise based on market conditions. They are deeply rooted in financial theory and are used for pricing MBS.
- **Key Research**: Explore studies that apply option-theoretic models to assess the economic value of prepayment options and their impact on MBS pricing.

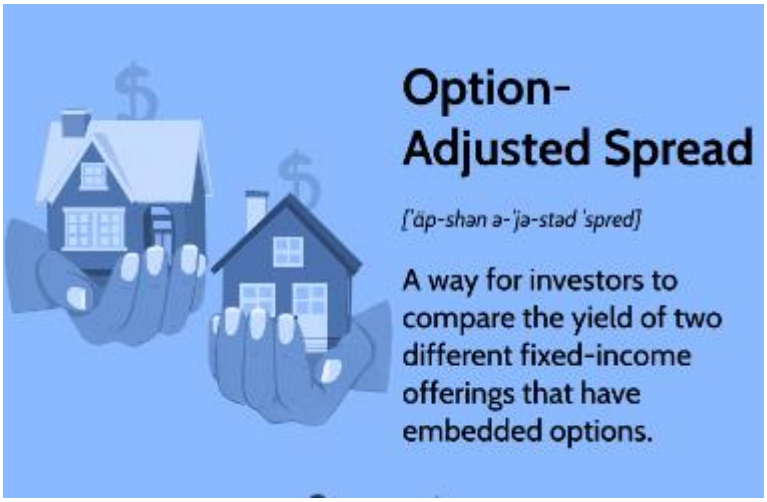
e. Econometric Models

- **Overview**: Econometric models, such as ARIMA (AutoRegressive Integrated Moving Average), are used to analyze time-series data on prepayment rates, allowing for forecasting based on historical trends.
- **Key Research**: Find studies that use econometric approaches to model prepayment behavior, particularly those that account for macroeconomic variables.

5. Association with Prepayment

i) Research how prepayment risk is typically measured and incorporated into models.

This could involve metrics like option-adjusted spread (OAS) :



a. Option-Adjusted Spread (OAS)

- **Overview**: The Option-Adjusted Spread is a key metric used to measure the risk of prepayment. OAS adjusts the yield spread of a bond (e.g., MBS) by accounting for the embedded prepayment option that borrowers have.
- **Calculation**: OAS is calculated using a model that simulates different interest rate scenarios and estimates the spread over the risk-free rate after adjusting for the likelihood of prepayments.
- **Interpretation**: A lower OAS indicates higher prepayment risk, as the investor is compensated less for taking on the risk of early prepayment. Conversely, a higher OAS suggests lower prepayment risk.

b. MBS Duration and Convexity

- **Duration**: Measures the sensitivity of the MBS's price to changes in interest rates. In the context of MBS, duration is adjusted to account for the likelihood of prepayments, which can shorten the expected life of the security.
    - **Negative Duration**: MBS often exhibit negative convexity, meaning that as interest rates decline (making refinancing more attractive), prepayments increase, shortening duration.
- **Convexity**: Refers to the rate of change of duration as interest rates change. In MBS, negative convexity occurs when prepayment speeds up as interest rates fall, leading to a decrease in the value of the MBS.

ii) Incorporating Prepayment Risk into Models

a. Option-Theoretic Models

- **Application**: These models treat prepayment as a financial option that borrowers can exercise, typically when interest rates fall. The prepayment option's value is estimated using techniques similar to those used for pricing financial derivatives.
- **Incorporation**: Prepayment risk is incorporated into the pricing of MBS by adjusting for the likelihood of prepayment in different interest rate scenarios. OAS is often used in these models to account for the risk-adjusted spread.

c. Regression Models

- **Application**: In a regression framework, prepayment risk can be modeled by including variables that influence prepayment behavior, such as interest rates, credit scores, and loan characteristics.
- **Incorporation**: The model may include the volatility of interest rates or economic conditions as predictors, capturing the uncertainty and risk of prepayments.

d. Machine Learning Models

- **Application**: Advanced machine learning techniques can be used to model prepayment risk by learning from historical data patterns.
- **Incorporation**: Feature importance methods can highlight which variables most influence prepayment behavior, and uncertainty measures (like prediction intervals) can be used to assess prepayment risk.

6.Summary of Potential Features for Prepayment Risk Modeling

Key Features:

1. **Interest Rates**
    - **Interest Rate Spread**: The difference between the original mortgage rate and current market rates. A wider spread increases the likelihood of prepayment as borrowers seek better rates.
2. **Borrower Characteristics**
    - **Credit Score**: Higher credit scores may indicate a higher likelihood of prepayment due to better refinancing opportunities.
3. **Loan Attributes**
    - **Loan-to-Value (LTV) Ratio**: Lower LTV ratios suggest more equity, which can lead to higher prepayment as borrowers may refinance more easily.
    - **Debt-to-Income (DTI) Ratio**: Higher DTI ratios may lead to higher prepayment risk if borrowers refinance to reduce monthly payments.
4. **Macroeconomic Conditions**
    - **Economic Growth**: Periods of economic growth generally lead to higher prepayments due to increased income and refinancing opportunities.
    - **Housing Market Conditions**: Rising home values can encourage prepayments through refinancing or selling.