

China RMBS FF Regression Model

Model Summary Document (MSD)

1. Model Purpose and Description

China RMBS FF logistic regression model is a criteria model with loan-by-loan mortgage data in China to address the risk relativity of each risk factor, while the base PD will be projected with market/aggregate data separately, due to data availability. The model estimates will act as an objective starting point for the China RMBS rating criteria assumptions.

The model is a multiple logistic regression that predicts default with risk factors including unindexed current loan-to-value, current debt-to-income, employment group and performing years.

2. Model Assumptions

The model estimates contribute to criteria assumptions in the “China Residential Mortgage Rating Criteria”. Statistical assumptions of logistic regression were diagnosed during initial development as well as 3-year validation testing. Diagnostic tests include statistical significance and multicollinearity.

3. Model Inputs

Three credit risk factors and one controlling factor, which can be derived from the data fields of the CCB’s data template, are used as the inputs:

- Unindexed Current Loan-To-Value (cLTV),
- Borrower Employment Group,
- Current Debt-To-Income (DTI) Group,
- Performing Year (i.e. the length of modelling window).

The variables are selected with stepwise procedure.

4. Model Development Data

The data received for initial model development are 11 RMBS pools of China Construction Bank (CCB), the largest RMBS issuer in China, and the latest pool cut date is April 2020. Because of the short RMBS history in China, the performing window is from 8 months to 5 years only. The total number of loans is 422,423 in the input dataset and 419,307 in the cleaned dataset for modelling. There are several data limitations in the current available data, including short performing history and inconsistent risk indicator for a small number of closed loans.

The data received for 3-year validation are two new CCB RMBS pools: one has initial pool-cut date of August 2020 and latest pool-cut of March 2024; the other has initial pool-cut of March 2021 and latest pool-cut of March 2023. The total number of loans is 84,346 in the input dataset and 83,874 in the cleaned dataset for testing. The data volume is relatively small compared to initial development and significant prepayment spikes are observed in Sep-Oct 2023 and Feb 2024 which may undermine default sensitivities.

5. Model Limitations

There are several data limitations in the data used for development and testing:

- Relatively short loan performing history,
- Limited data fields in the data template,
- Suspected unreliable data quality for some important information (e.g. income),
- Inconsistent risk indicator for a small number of outstanding loans and closed loans,
- One-off, policy-driven significant prepayment spikes in the new testing data, etc.

6. Material Changes

The current model version is 1.0.0. Any future material changes will be disclosed in this section.

Contact Information

model_disclosure@fitchratings.com