Behavioural modelling – prepayment of term loans

Anirban Naskar, Genpact

Arunima Banerjee, Genpact

2017

Table of contents

The views expressed here are those of the authors and do not necessarily represent or reflect the views of their employers.

Abstract

Banks for International Settlement (BIS) and various banking regulators mandate banks to factor prepayments for fixed interest rate loans for assessment of liquidity risk and interest rate risk in banking book. There are various methods of modelling prepayments of such loans. Broadly, these methods are categorised under static model, the model which does not consider financial incentive of prepayment, and the dynamic model, the model which considers financial incentive of prepayment. In this paper, we have simulated loan data based on various real-life factors and used the data to demonstrate various static and dynamic modelling approaches.

Acknowledgements

We would like to thank Kaushik Das, a risk professional, and, our colleagues from Genpact, especially Sidharth Reddy and Sadanand Tutakne. We would also like to thank many other industry experts; with whom we have discussed this subject at various occasions.

Key-Words: prepayment, foreclosure, behavioural modelling, term loan modelling, interest rate risk in banking book, IRRBB, liquidity risk, asset and liability management, balance sheet management, loan book management, risk management

1. INTRODUCTION

The core activities of a bank are to raise funds through deposits and market borrowings, and deploy the same through loans and advances and various investments. Banks tend to take advantage of upward sloping yield curve by sourcing funds in short term and deploying these funds in long term. For some of the assets and liabilities, the actual maturities vary from the contractual maturities. One such item is term loans. Although terms loans are disbursed with fixed tenure, borrowers have the options to voluntarily prepay, in full or part. Customers tend to prepay when they have surplus liquidity (their equity) which can be used to reduce their debt burden, hence, interest expenses. Sometimes borrowers can refinance the loans from other sources at lower rate of interest. Such interest benefit, net of prepayment penalty, can trigger prepayment events. From lenders' perspective, prepayment assumptions impact earnings, valuations and risk management planning. For example, during prepayment process, treasury managers are left with additional cash balance than they would have anticipated. This additional cash can be deployed in economic usages, thus increasing revenues or reducing expenses. Reinvestment risk, earnings volatility, valuation volatility and liquidity risk are some of the risks which are resultant of prepayment.

In such a scenario, measurement of prepayment risk becomes important, and, is also required for regulatory compliance in banks. It is important to capture prepayment risks in interest rate risk from banking book (IRRBB) and liquidity risk measurements, along with other regulatory requirements. In the latest IRRBB guideline[1] from Basel Committee, the Committee expected banks to follow nine principles for managing IRRBB. The principle number five expects, "In measuring IRRBB, key behavioural and modelling assumptions should be fully understood, conceptually sound and documented. Such assumptions should be rigorously tested and aligned with the bank's business strategies." In the guideline, it also mentioned the Committee's expectations around the implementation of prepayments. Banks must determine or supervisors must prescribe the baseline prepayment assumptions, that is, conditional prepayment rate (CPR). This must be assessed differently for loans in different currencies. We also expect prepayments to vary with macro-economic scenarios that are required for the purpose of scenario analysis and stress testing. Similarly, the liquidity risk guideline[2] from Basel Committee stated, "A bank should have a robust liquidity risk management framework providing prospective, dynamic cash flow forecasts that include assumptions on the likely behavioural

[1] Standards, Interest rate risk in the banking book, Bank for International Settlement, April 2016
[2] Principles for Sound Liquidity Risk Management and Supervision, Bank for International Settlement September 2008

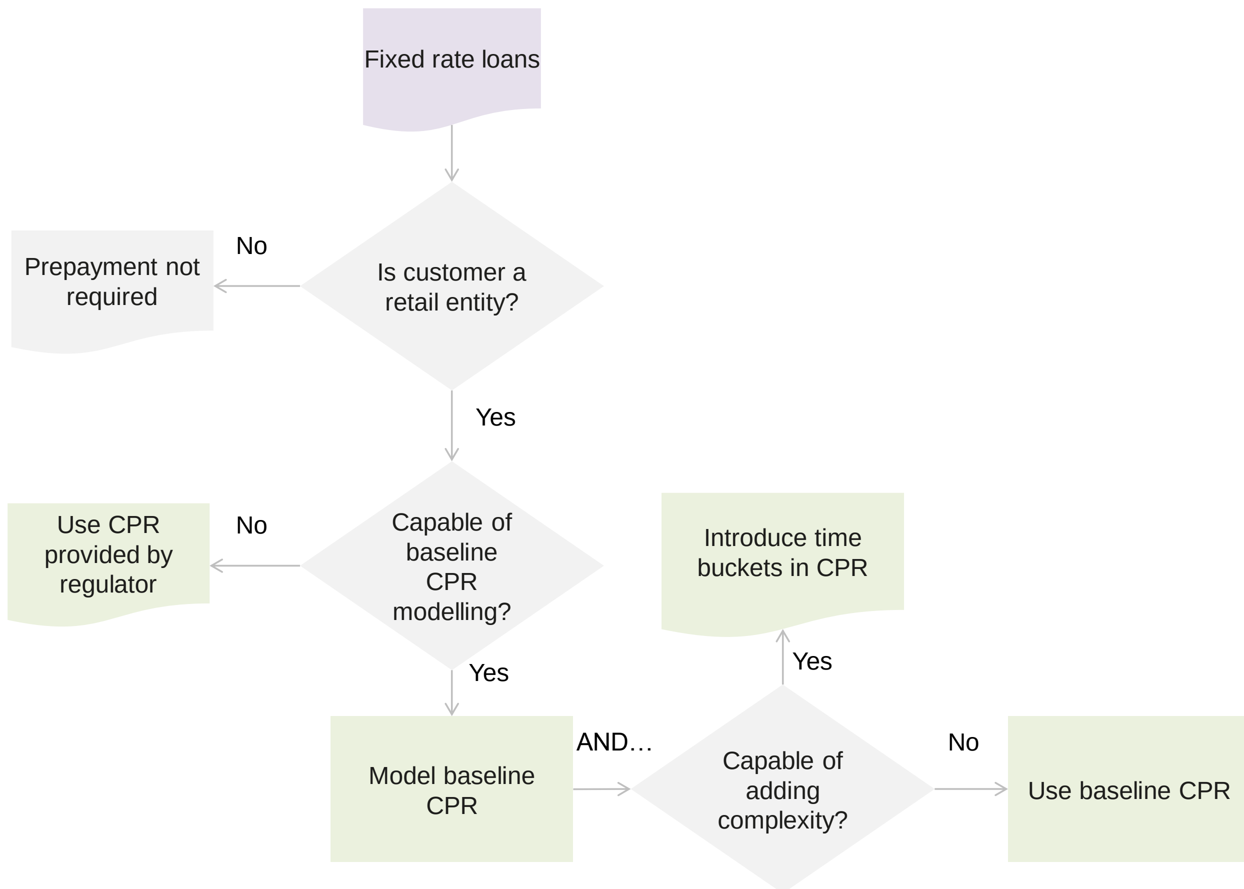
responses of key counterparties to changes in conditions and are carried out at a sufficiently granular level." Accordingly, banks are to measure impact of prepayments in forecasting the future cash flows for term loans.

As banking regulators of various countries customise the BCBS's guidelines to incorporate local market practices and nature of risks, regulations on IRRBB and liquidity risks mandate incorporation of prepayment risks in various risk metrics, and ultimately, in regulatory capital calculations. In the USA, the Office of the Comptroller of Currency (OCC), at its liquidity booklet[3] stated, "Effective management and control of the liquidity risk stemming from funding gaps depends heavily on the use of operational cash flow projections and the reasonableness and accuracy of the assumptions that are applied." It also mentioned that "Highly volatile or unpredictable asset amortization (prepayments)" is one of the many factors that affect the cash flows. In the context of assessing risk of securitised products, the Fed stated[4], "The prepayment of assets underlying ABS may create prepayment risk for an investor in ABS. Prepayment risk may not be adequately reflected in agency ratings of ABS. Examiners should determine that a banking organization investing in ABS has analysed the prepayment risk of ABS issues in its portfolio." It continues, "Prepayment risk for ABS should be incorporated into an organization's "net income at risk" model if such a model is used." In fact, for some of the credit risk models, the life-time recoverable include expected prepayments by borrowers.

The following figure summarises the Basel IRRBB guidelines pertaining to the treatment of prepayment.

[3] Comptroller's Handbook, Safety and Soundness, Liquidity, June 2012, OCC
[4] Examination Guidelines for Asset Securitization, SR letter 9016a1, Federal Reserve System

Exhibit 1: Basel guidelines for treatment of prepayments



This paper intends to address these regulatory requirements through appropriate prepayment modelling techniques. We have studied various prepayment models and summarised our findings in section 2 of this paper. Subsequently, we have narrowed down some of the approaches which would comply with regulatory requirements for banks with varied size and complexity. Those approaches will be provided in section 3. We have demonstrated how the selected methodologies can be applied on the simulated data. The simulation procedure will be given in section 3.2. The application methodology and the results will be produced in section 3.3. We will present our concluding remarks in section 4.

The intention behind this paper is to demonstrate some of the regulatory-compliant methodologies on a given dataset. The specific results, produced here, may not be relevant for a given bank, as the nature of loan books will vary from bank to bank. However, the methodologies will be useful for developing prepayment models which will be compliant with regulatory expectations.

2. LITERATURE SURVEY

Under the stated objectives in mind, we have studied various prepayment modelling methodologies. Some of these methods are obsolete and some are presently being used in industry. Prepayment models are broadly of two types, one in which refinances incentives are considered, and the other in which refinances incentives are not considered. The former kind is called ‘dynamic model’ and the later ‘static model’.

2.1 Static Prepayment Models

Static prepayment approach, in spite of its limitations, is still considered as a simple approach to describe prepayments. We have explored the following four common static models, some of these are used in the industry and some are only referred to for historical significance. Some of the notable dynamic models for prepayment are described below.

2.1.1 12-year life

This is one of the earliest static models built during the 1970s. The approach[5] assumes that there is no prepayment for first 12 years and then the entire outstanding is prepaid. The major shortcoming of this model is that the mortgage market has experienced regime shifts since that period, and, average behaviour of mortgage has changed since then. This model does not consider various factors like loan characteristics, prepayment incentives, etc. Accordingly, this model is too simplistic to be used in current period.

2.1.2 FHA Experience

This is another model which is no more in use but has historical significance. This model was developed using data maintained by Federal Housing Administration (FHA), USA since 1930s. According to Lawrence Rosen[6], the model was built using FHA data on 30-year FHA insured mortgages and is derived from a probability of survival table of the FHA. For the first time, seasoning (age) of loans was introduced in prepayment modelling. Frank J. Fabozzi[7] observed

---

[5] Philippe Priaulet and Lionel Martellini, “Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies”, Page 601, Wiley publications, May 2003

[6] Lawrence Rosen, “The McGraw-Hill Handbook of Interest, Yields, and Returns”, page 127, McGraw-Hill, 1995

[7] Frank J. Fabozzi, “Collateralized Mortgage Obligations: Structures and Analysis”, page 22, Wiley publications, April 1993

that the since prepayment rates are closely linked to interest rate cycles, using the average prepayment rates over various cycles as estimates for prepayments will not be effective. Also, since FHA tables were published periodically, there was ambiguity in terms of the identifying the correct FHA table that should be used in the calculation of prepayments. Also, the model was developed using mortgage data and hence, this could not be used for other term loans, like auto loans.

2.1.3 CPR and PSA

Conditional prepayment rate (CPR), also referred to as constant prepayment rate, model is most widely used prepayment model for regulatory compliance, especially for IRRBB and liquidity risk compliance. CPR is an annualised of prepayment compared to principal outstanding. CPR is defined as,

$$CPR = 1 - (1 - SMM)^{12}, \qquad where, single\ month\ mortality\ (SMM)\ is\ defined\ as,$$

$$SMM = \frac{Expected\ principal\ prepayment\ in\ one\ month}{Outstanding\ principal\ at\ the\ end\ of\ the\ month}$$

Under Public Securities Association (PSA), the major participants in the mortgage securities market had agreed[8] on a standardised method to calculate yield of collateralised mortgage obligations. This model eliminated the confusion that arose out of nonstandard prepayment assumptions made by different market participants. In this model, the rate of prepayments (for mortgages) would linearly increase from 0% to 6% in 30 months, that is, at the rate of 0.2% per month, and then be constant at 6%. This base assumption is also referred to as '100% PSA'. The prepayment assumptions can be scaled up / down linearly. For example, '200% PSA' means CPR increases from 0% to 12% in 30 months, which is, at the rate of 0.4% per month, and then be constant at 12%.

2.2 Dynamic Prepayment Models

The dynamic prepayment models have increasingly gained importance as these models are effective for cash flow modelling and also capture the sensitivity of prepayments to market interest

[8] The New York Times, "Calculation Standard Set on C.M.O. Yields", June 14, 1985

rate, thus dynamic in nature. Most of the dynamic models consider refinancing incentive, seasoning, seasonality and burnout effect[9] to be model factors. Some of the notable dynamic models for prepayment are described below.

2.2.1 Andrew Davidson & Co. (ADCO) Model

The ADCO Model considers the turnover and seasonality, refinance incentive, cash out effect and credit cure effect as the factors that primarily guide prepayment.

Refinance incentive is the most important factor driving prepayments and is primarily driven by the level of the current mortgage rate relative to the weighted average gross coupon (GWAC) of the pool. Higher the difference between these two rates higher is the incentive to prepay.

Turnover and seasonality is considered to be the another important driver of prepayment and it tends to be seasonal in nature with turnovers increasing during the summer.

Cash out effect considers the steady appreciation of home value in the near past and implies the equity building up in the home. This factor depends on current ageing of the loan, refinancing incentive and the magnitude of home price appreciation. This factor can cause prepayments even by borrowers having no interest rate refinancing incentive.

The credit cure effect takes into account the spread of borrower ratings before and after credit curing. A greater spread[10] implies a greater likelihood to refinance post credit curing.

---

[9] "A period of slowing mortgage prepayment within a mortgage backed security (MBS). This usually occurs after the mortgages start to mature. When some percentages of the underlying loans fail to prepay after an interest rate cycle, this is known as burnout. Those borrowers who did not refinance during the first interest rate cycle are less like to do so if interest rates drop again.", Investopedia LLC., http://www.investopedia.com/

[10] When borrowers have poor credit worthiness, the borrowing rate is considerably higher than the prevailing market rate. However, with improvements in their credit quality, these borrowers become eligible for lower rates which are closer to the prevailing market rates, thereby increasing the probability of refinancing.

2.2.2 Bloomberg Prepayment Model (BPM)

The BPM approach considers three independent components of which housing turnover and refinancing components are similar to those in the ADCO method explained in section2.2.1. The third component is the default and curtailment component unique to the BPM methodology. Default model captures the prepayment that occurs due to sell off of the property post a loan defaults. Curtailment shortens the maturity of the loans and reduces the WAL of the pool. Full payoff (foreclosure) normally occurs during last few years of the tenure.

Apart from the components cited above, BPM also incorporates the loan level components like loan size, credit score, loan to value ratio and occupancy type as instrumental variables in modelling prepayments.

2.2.3 Solomon-Smith-Barney (SSB) Model[11] [12] [13]

Similar to the ADCO model, the SSB model uses refinance incentive and housing turnover as important components affecting prepayment. Additionally, the model uses two other components, a defaulter model and a curtailment and payoff model, which is similar to those in the BPM model.

2.2.4 Wharton Model

In 1992, Zenios and Kang came up with the Wharton prepayment model[14]. Though, this was a mortgage prepayment model, it could be useful for other term loans after customisations. The model considered, refinancing incentive, seasonality effect, ageing effect or seasoning, and burnout effect.

---

[11] First published in Salomon Smith Barney's proprietary analytical system, The Yield Book, in August 2000, this is a fully loan level prepayment model

[12] Lakhbir Hayre, "Salomon Smith Barney Guide to Mortgage-Backed and Asset-Backed Securities", page 549, Wiley publishers, April 2001

[13] Lakhbir Hayre and Arvind Rajan, "Anatomy of Prepayments: The Salomon Brothers Prepayment Model", Salomon Brothers, June 1995

[14] Kang and Zenios, "Complete prepayment models for mortgage backed securities", 1992

2.2.5 Goldman-Sachs (Richard-Roll) Model

This model is very similar to the Wharton model. The CPR is determined as a function of market mortgage rate and contract rate, mortgagor costs, age of the loan, month of the year and the interaction between these variables. The seasonality and burnout factor are derived as a function of age of the loan and the refinancing incentive. The refinancing incentive is derived as a function of the ratio of contract rate and market mortgage rate.

3. ANALYSIS

3.1 Approach

Since the CPR prepayment methodology is the most widely used prepayment model for regulatory compliance, especially for IRRBB and liquidity risk compliance, we focused on this methodology, for our analysis, in this paper. However, based on the systems and infrastructure available, banks can adopt any of the other techniques of prepayment modelling identified above. We use simulated portfolio data to model prepayment rates. Though simulations have been used here, we incorporated real life factors that affect prepayments and cash flows. Therefore the simulated scenarios used here are similar to actual real life scenarios. We have provided details of the simulation process in section 3.2.

We used the simulated data to fit a model that can be used to predict CPR numbers. The methodologies followed for model fitting include one-step and two-step regression[15] models. Moreover, depending on the explanatory factors included in the model, we used static and dynamic approaches where the static model considers seasoning as the only explanatory variable and the dynamic model considers refinance incentive and seasonality along with seasoning as explanatory variables. The details of the modelling approach and the alternate models explored have been provided in section 3.3.

3.2 Data simulation for modelling

We simulate a portfolio of 1000 term loans. These loans are between 15 to 20 years. In the portfolio, 50% of the loans have a flexible tenure (a prepayment alters the loan maturity while keeping the EMI unchanged) and 50% of the loans have a flexible EMI (a prepayment alters the loan EMI while keeping the loan tenure unchanged). For each loan, we start with a loan amount, term and start date. For this simulation, we freeze the portfolio at certain date and observe its behaviour for the analysis period. This also ensures than we do not consider incremental book. The details of the simulation parameters are provided in Exhibit 2, Exhibit 3 and Exhibit 4.

Exhibit 2: Prepayment probabilities

[15] Various kinds of linear regressions have been used, algorithm minimizes sum of squared errors in all cases.

Loan rate - prevailing rate *		< 1 %	1 – 3%	3 – 6%	> 6%
Seasoning	< 5 years	Zero	Low	Low	Medium
	> 5 years			Medium	High

** Simulated from 10Y GSEC + Spread*

Exhibit 3: Consequence of prepayment

Outcome	Associated probability
Maturity remains constant, EMI reduced	50%
EMI remains constant, maturity advanced	50%

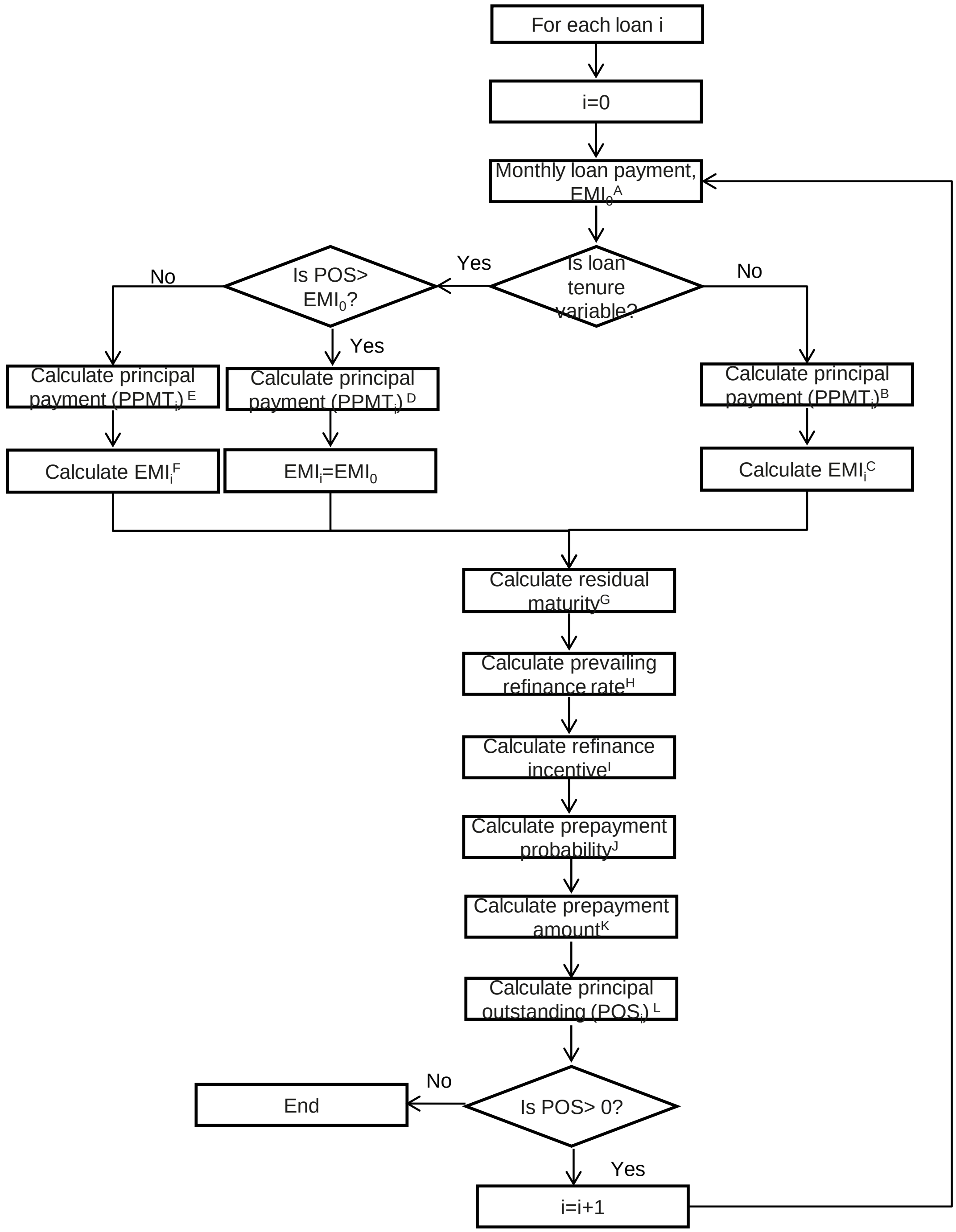
Exhibit 4: Amount of prepayment

Outstanding / loan amount	100-10%	10-0%
Low*	0 to 25%	25 to 50%
Medium	25 to 50%	50 to 90%
High	50 to 90%	90 to 100%

**As derived from Exhibit 2*

Given these, we simulate the cash flows for each loan as described below.

Exhibit 5: Procedure to simulate cash flows



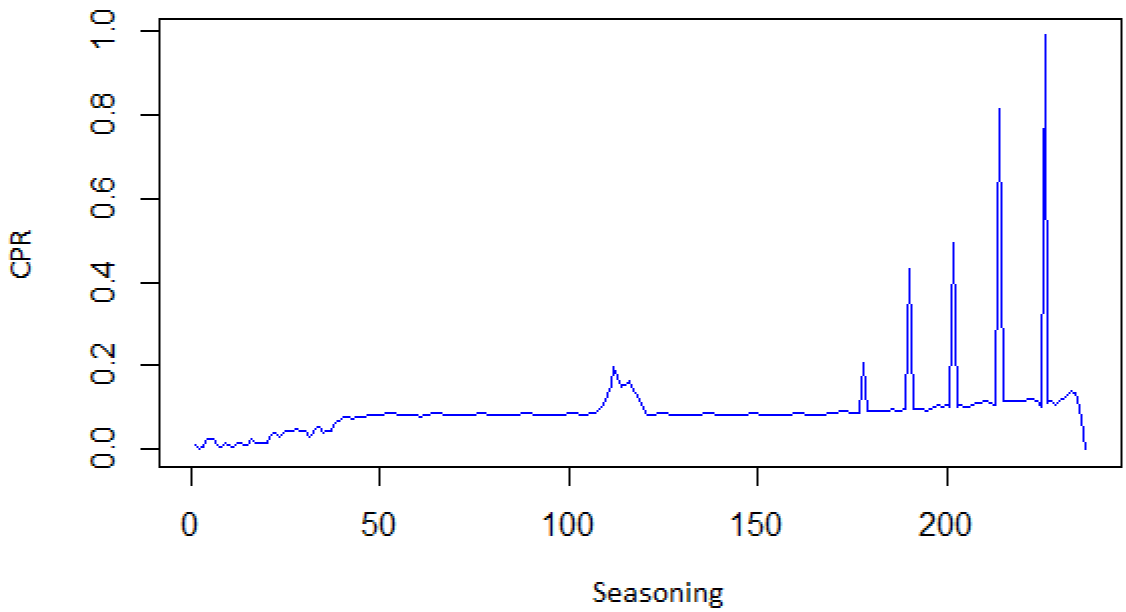
3.3 Modelling methodology and results

3.3.1 Data treatments

The objective is to fit a model using the simulated data. We begin with cleaning the data to remove/replace observations that are not likely to get repeated in future.

The following figure shows the CPR rates from the simulated portfolio.

Exhibit 6: Actual CPR from loan portfolio (for 237 months)

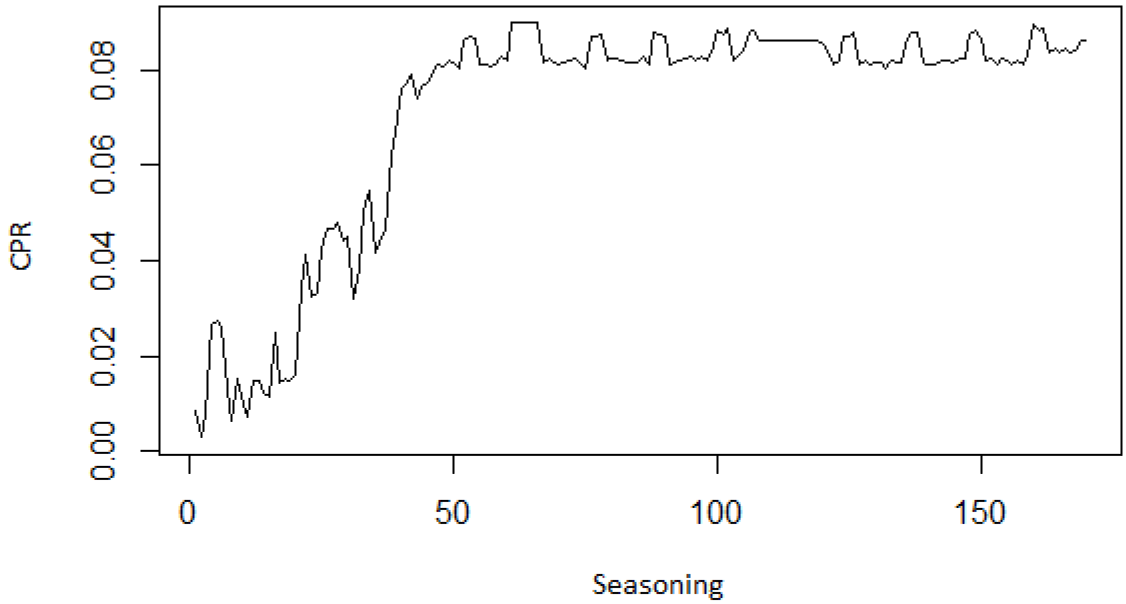


We observe from the above figure that the CPR numbers beyond the 170th observation are not in line with the general CPR behaviour. This occurs due to outstanding balances that are low due to previous prepayments and loan closures. Thus, we consider observations till the 170th period to model the CPR rates.

Moreover, we see that there is a spike between observations 100 and 125. This is due to the underlying 10 year US GSEC yields around this period. This observation is also outside the general behaviour of CPR rates and thus, we cap the values around this period to the 90th percentile value of observations 51 to 170.

The final dataset of CPR numbers that is used for modelling purposes is given in the following figure.

Exhibit 7: After outlier treatment (170 months)



The following table also summarizes the descriptive statistics of the final CPR rates series.

Exhibit 8: Descriptive statistics (in %)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev
0.30%	7.71%	8.19%	7.12%	8.60%	8.97%	2.45%

3.3.2 CPR Modelling

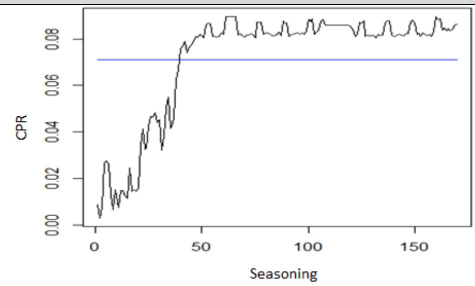
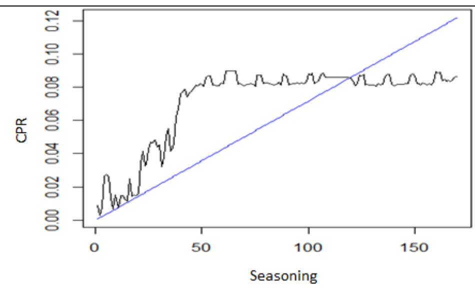
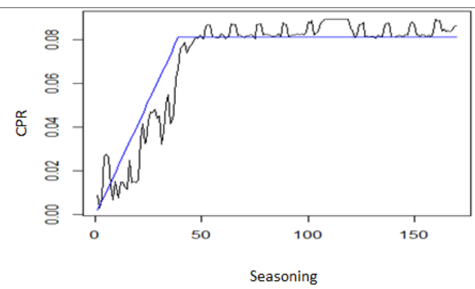
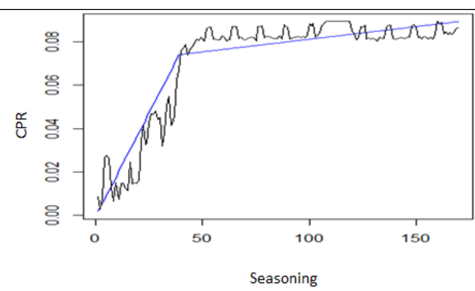
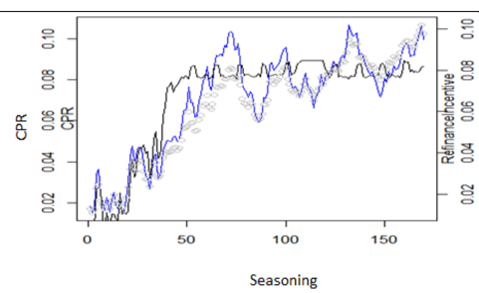
We model the CPR series using static and dynamic approaches. We use ordinary least squares (OLS) regression to model the CPR numbers as a function of different explanatory variables.

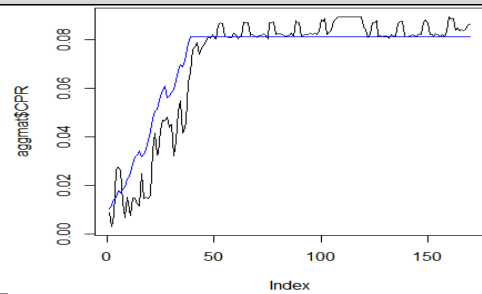
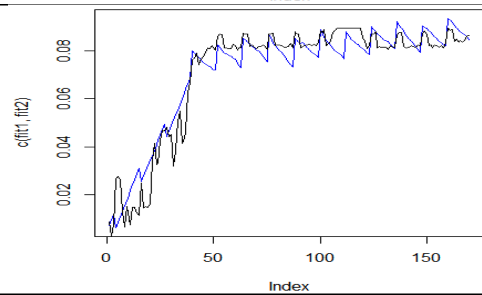
We calculate the model forecasting error as below:

$$error = \sqrt{\frac{\sum_{i=1}^{n}(Actual\ CPR - Predicted\ CPR)^2}{n}}$$

We have modelled CPR using both, static and dynamic CPR modelling methods. Here, we summarize the results obtained, followed by descriptions of each model type.

Exhibit 9: Summary of results from alternate methods

Sl. No.	Model category	Model Type	Forecast trends	S.S.E.
1	Static	One CPR for all bucket		2.45%
2		One step linear model		2.60%
3		Two step linear model where CPR increases from zero and stabilises beyond a point		0.91%
4		Two step linear model where CPR increases from zero till a point beyond which its growth is minimal		0.80%
5	Dynamic	One step linear model		1.24%

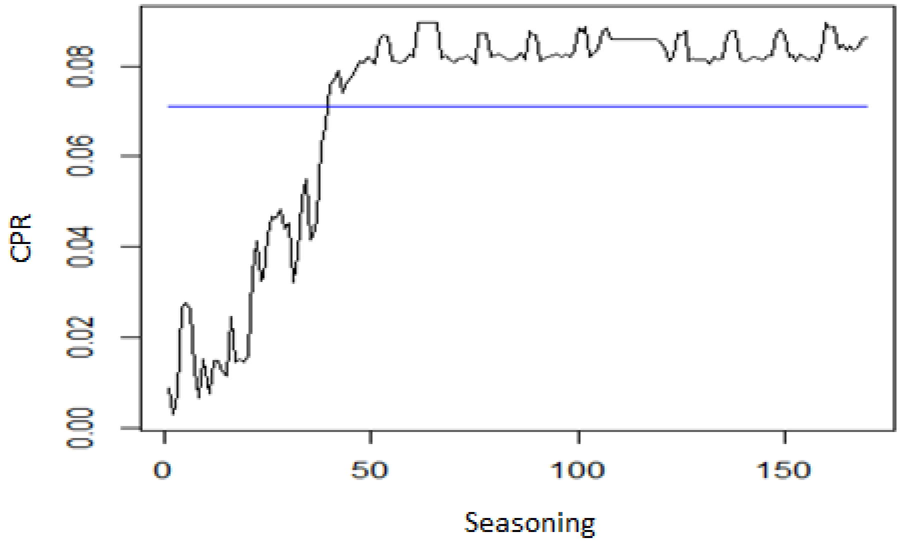
Sl. No.	Model category	Model Type	Forecast trends	S.S.E.
6		Two step linear model where CPR increases from zero and stabilises beyond a point		0.89%
7		Two step linear model where CPR increases from zero till a point beyond which its growth is minimal		0.69%

Static CPR Modelling

3.3.2.1.1 One CPR for all buckets

In this approach, we determine a constant CPR number across all maturity buckets. This is calculated as the simple average of the CPR series and results into a CPR number of 7.12%. For this model, we get a prediction error of 2.45%.

Exhibit 10: Static CPR Model – One CPR for all buckets

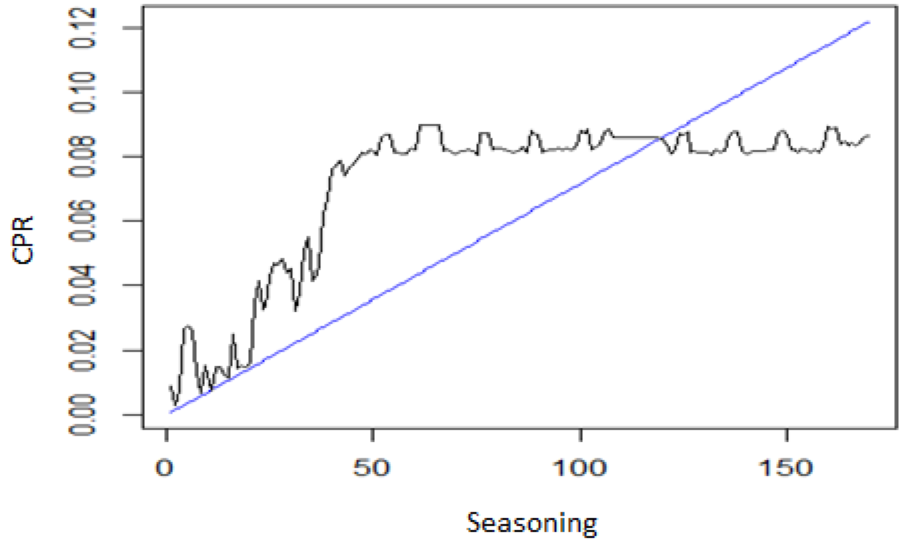


However, we see a constant CPR number, as given by the blue line, across all maturity buckets that averages out the actual behaviour of CPR curve over time, and is particularly erroneous during the initial periods of origination. Hence, we fit a linear model to the CPR series using ordinary least squares approach.

3.3.2.1.2 One step linear model

In this approach, we use a linear model to predict CPR numbers over different levels of seasoning or months into the loan. Thus, seasoning becomes the only explanatory variable in our model. We use a simple linear regression model for this purpose. This model generates a prediction error of 2.60%. The following figure shows the plot of the actual vs predicted CPR numbers.

Exhibit 11: Static CPR Model – One-step CPR model



In the above figure, positively sloped linear curve fit reflects the actual pattern in CPR numbers during the initial stages of seasoning, the actual CPR numbers flatten out over time and follows a constant trend after a point of time. This cannot be captured using a one-step linear regression model and hence we move towards a two-step regression model.

3.3.2.1.3 Two step linear model

For this approach, we split the actual CPR numbers series in two segments - early and seasoned and fit two separate models to each segment. However, since the CPR rates are continuous in nature, we forcibly ensured that the two segments meet at one point instead of being completely discrete segments.

To identify the breakpoint, we identified the optimal point which minimizes the total squared error in model prediction as obtained using the fitted models for the two segments.

While for the first segment, we used a positively sloped linear model with no constant, for the second segment we have used alternate assumptions as below:

- Linear segment with zero slope, that is, constant
- Positively sloped linear model with non-zero constant, slope is different from that of first segment

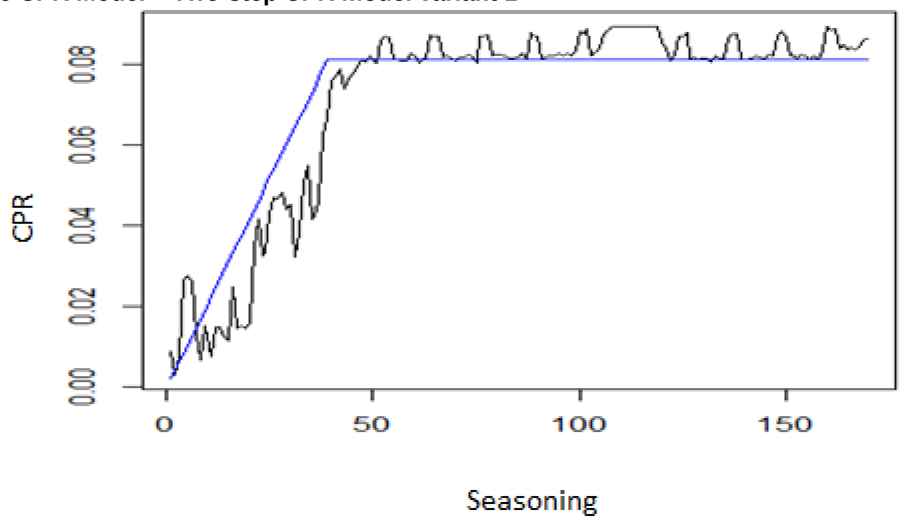
3.3.2.1.3.1 Linear segment with zero constant

Using this approach, we fit the following models to the two segments:

- Segment 1: $y = b * seasoning$
- Segment 2: $y = a$

Using this approach, the 37th observation was identified as the break point. This minimised the model error and the model prediction error came down to 0.91%.

Exhibit 12: Static CPR Model – Two-step CPR model variant 1



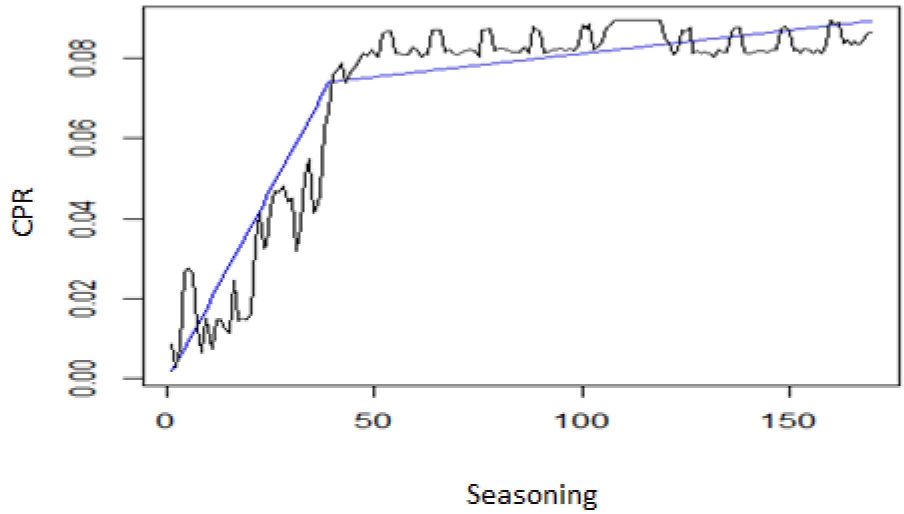
3.3.2.1.3.2 Linear segment with non-zero constant

Using this approach, we fitted the following models to the two segments:

- Segment 1: $y = b_1 * seasoning$
- Segment 2: $y = a_2 + b_2 * seasoning$

Using this approach, the 37th observation was identified as the break point which minimised the model error and the model prediction error came down to 0.80%.

Exhibit 13: Static CPR Model – Two-step CPR model variant 2



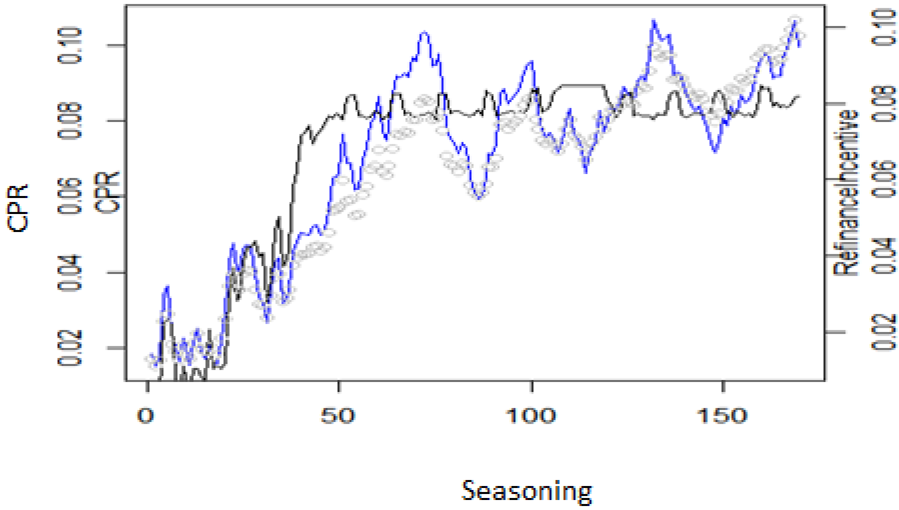
Dynamic CPR Modelling

For the dynamic CPR modelling, we used the same model types as used for static CPR modelling and also allowed for inclusion of other explanatory variables including refinance incentive and seasonality.

3.3.2.1.4 One step linear model

In this approach, we use a linear model to predict CPR numbers as a function of seasoning, seasonality and refinance incentive. We use a simple linear regression model for this purpose. This model generates a prediction error of 1.24%. The following figure shows the plot of the actual vs predicted CPR numbers.

Exhibit 14: Dynamic CPR Model – One-step CPR model



3.3.2.1.5 Two step linear model

We follow a similar approach to develop a dynamic two-step model as followed for the static two-step model.

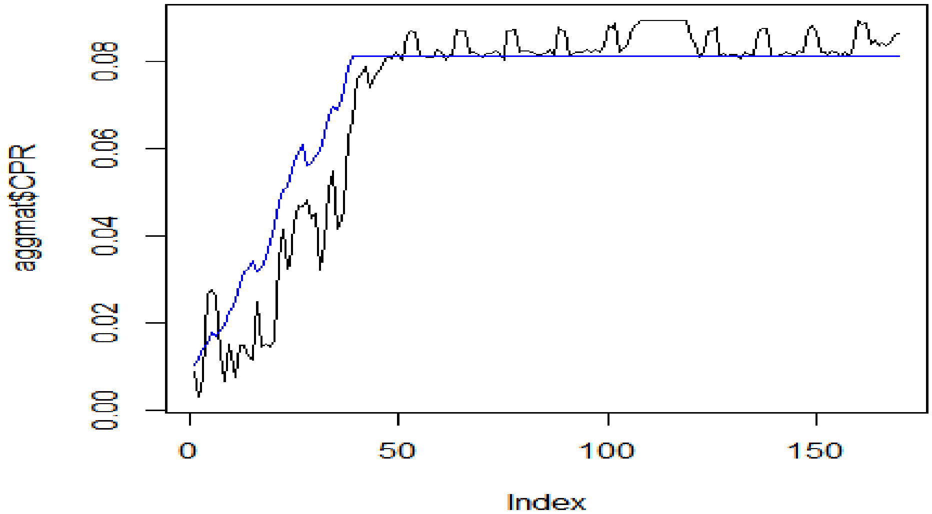
3.3.2.1.5.1 Linear segment with zero constant

Using this approach, we fit the following models to the two segments:

- Segment 1: $y = b * seasoning + c * refinance\ incentive + d * seasonality$
- Segment 2: $y = a$

Using this approach, we identified the 37th observation as the break point. This minimised the model error and the model prediction error came down to 0.89%.

Exhibit 15: Dynamic CPR Model – Two-step CPR model variant 1



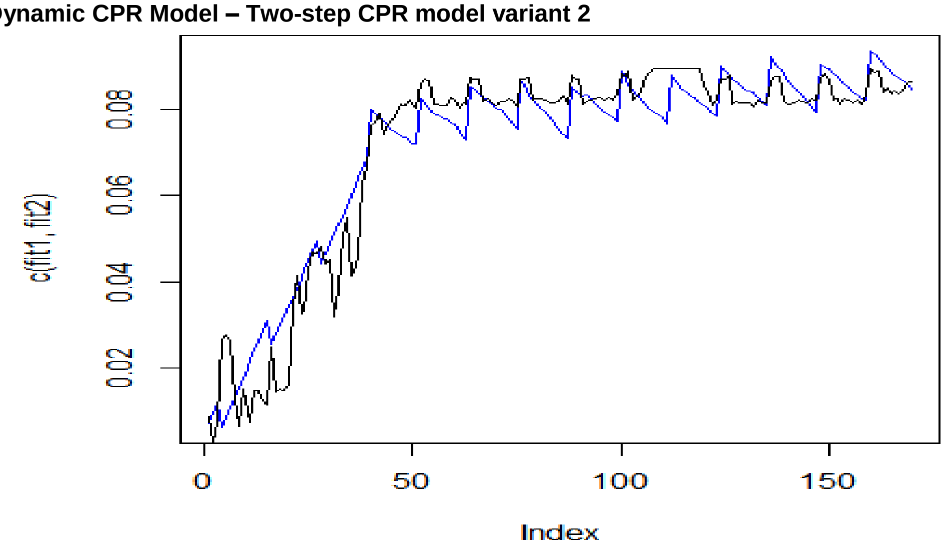
3.3.2.1.5.2 Linear segment with non-zero constant

Using this approach, we fit the following models to the two segments:

- Segment 1: $y = b_1 * seasoning + c_1 * refinance\ incentive + d_1 * seasonality$
- Segment 2: $y = a_2 + b_2 * seasoning + c_2 * refinance\ incentive + d_2 * seasonality$

Using this approach, we identified the 37th observation as the break point. This minimised the model error and the model prediction error came down to 0.69%.

Exhibit 16: Dynamic CPR Model – Two-step CPR model variant 2



4. CONCLUDING REMARKS

A wide variety of modelling techniques can be applied to model prepayment. The prepayment modelling approach, used by an entity, depends on specific features of the portfolio. Hence no single approach can be recommended to be optimal. Therefore, the identification of the final modelling approach needs to take into account the portfolio structure and model performance using the identified methodology. Validation and monitoring are thus crucial to such models.

Many other factors, which are not covered in this paper, also affect prepayment trends. Such factors cannot be generalized and should be given due weightages. For example, we should understand how loans are being marketed, sanctioned, priced and serviced while modelling. Subprime mortgage markets grew after 1980 and collapsed during the 2008 crises. The subprime products were structured and marketed differently than the prime products. Subprime customers borrowed despite high prepayment penalties which created a barrier for refinancing. Subsequently, post crises, the prepayment scenario was completely transformed under the new interest rate regime where the rates were kept low by the central bank. Despite low interest rates, banks had to struggle with defaults rather than prepayments which should have been 'natural' phenomenon of high prepayment in low interest rate periods.

ABOUT THE AUTHORS

Anirban Naskar is a Senior Manager in Genpact. He works in asset and liability management (ALM) and market risk management for banks and consulting organizations. He holds a Bachelor of Technology (with honours) as well as a Master of Technology from the Indian Institute of Technology, Kharagpur, India, and received his Master of Management (MBA equivalent) from the Indian Institute of Technology, Bombay, India. He is also a National Talent Scholar, awarded by the government of India.

Arunima Banerjee is an Assistant Manager in Genpact. She works in predictive statistical and mathematical modelling in the area of risk management. By education, she is a Master of Arts (Economics) from University of Delhi, India with specialisation in Game Theory and Econometrics.

Anirban Naskar
E: anirban.naskar@genpact.com

Arunima Banerjee
E: arunima.banerjee@genpact.com