|  |
| --- |
|  |
| TFRS9 Model Development Document |
| Housing Loan Portfolio |
|  |

Document History

This section documented the revision history and version control of this document. It shall record every major and minor revision of the model development document regarding the PD, EAD, LGD and SICR criteria models which are used for the purpose of calculation of ECL of the TFRS9 accounting book.

Revision History

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Version** | **Phase** | **Revision Date** | **Summary of Changes** | **Page** | **File name** | **Changed by** |
| 1.0 | 3.2 | 14-Jun-19 | Initial Version |  | KBank\_TFRS9\_Model\_Development\_Document\_HL\_V1.0\_190614 | Lalita L. |

Reviewer

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Version** | **Signature** | **Review Date** |
| Nittha Praechanya | 1.0 |  |  |
| Thepdanai Danswasvong | 1.0 |  |  |
| Panrit Tosukhowong | 1.0 |  |  |
| Nareerut Poolpun | 1.0 |  |  |

Approvals

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Version** | **Signature** | **Approve Date** |
| Duangporn Kit-o-pas | 1.0 |  |  |
| Sanphet Sukhapesna | 1.0 |  |  |

# Introduction

Probability of default (PD), Exposure at default (EAD), and Loss Given Default (LGD) are three components which are used to calculate credit risk capital and provision. In this document, we are focusing on Staff loan portfolio which is a Term Loan. The portfolio size of Other Unsecured products is about 287,583,879,070.17 Baht

In the current regulatory setting, the provisioning of expected credit loss is calculated from inherent risk parameters; the probability of default (PD), exposure at default (EAD) and loss given default (LGD). Each parameter is estimated by average through the economic cycle from 2007 to 2009 which can be considered as a downturn situation. However, these risk parameters does not reflect future risk exposure. Hence, the TFRS9 which is a new accounting standard is introduced. The impairment under TFRS9 setting also covers forward-looking components which should help improve financial stability and improve bank credit risk.

This document outlines the development process of all model related to the TFRS9 calculation. For each risk component, this document shall clearly state the scope (model usage), methodology considered, model development approach, final model and the initial validation results.

# TFRS9 Expected Credit Loss

On July 24, 2014, the International Accounting Standards Board (IASB) issued the final version of the ‘International Financial Reporting Standard (IFRS) 9 – Financial Instruments’. As a primary component of the new accounting standard, the IASB introduced a forward looking impairment model. The IASB thereby reacted to delayed recognition of credit losses identified as a weakness of existing accounting standards during the course of the global financial crisis (of 2007/08). In particular, the biggest critique of incurred loss approach under IAS 39 was the recognition of credit losses only upon evidence of a trigger event. In this regard, IASB’s approach of forward looking credit loss estimation was evident from the below extract.

*“The new standard requires an entity to recognise expected credit losses at all times and to update the amount of expected credit losses recognised at each reporting date to reflect changes in the credit risk of financial instruments. This model is forward-looking and it eliminates the threshold for the recognition of expected credit losses, so that it is no longer necessary for a trigger event to have occurred before credit losses are recognised. Consequently, more timely information is required to be provided about expected credit losses.” [“Project summary – IFRS 9 Financial Instruments”, IFRS Foundation, 07/2014, p.14][[1]](#footnote-1)*

Following the publication of IFRS 9 Financial Instruments in July 2014, the Basel Committee on Banking Supervision issued their ‘Guidance on Credit Risk and Accounting for Expected Credit Losses’ (GCRAECL) in December 2015. This covers in particular the impairment (Expected Credit Losses) element and how it should be embedded in and supported by internal processes.

Thai Accounting Standards are substantially converged with IFRS Standards, though the financial instruments Standards that are part of IFRS Standards have not yet been adopted. Thai Accounting Standards include several national financial instruments standards that differ from IFRS Standards. Henceforth TFRS9 can be considered as an adaptation from IFRS9 and replaces the existing TAS101.

***Principle 5 states****–A bank should have policies and procedures in place to appropriately validate models used to assess and measure expected credit losses. This presentation will provide an overview of the scope of work and the proposed validation approach for KBank, based on further discussions we will provide a more detailed view of the approach based on the complexity and materiality of the underlying models.*

Changes due to ‘*IFRS 9 – Financial Instruments*’ can be grouped into three categories.

* ***Classification and measurement***: Classification determines how financial assets and liabilities are accounted for in financial statements and, in particular, how they are measured on an ongoing basis:
  + Assets: one classification approach
  + Liabilities: addressing the volatility in profit or loss caused by changes in the credit risk of financial liabilities that are measured at fair value
* ***Impairments***: Forward-looking impairment model based on expected losses:
  + The new model requires entities to recognise expected credit losses at all times (12-month or lifetime expected loss) which includes measurement of changes in expected credit losses
  + It is no longer necessary for a trigger event to have occurred before credit losses are recognised
  + The new model is also accompanied by improved disclosures about expected credit losses and credit risk
* ***Hedge accounting***: Clear alignment with risk management:
  + The rules allow components of non-financial items to be hedged (previously not allowed by IAS 39)
  + IFRS 9 eliminates the distinction between financial and non-financial items and looks at whether a risk component can be identified and measured and therefore reflected in management activities

The primary change from IAS 39 to IFRS 9 is the evolution from an incurred loss view to a forward looking expected loss view which needs to be accounted for in the impairment models.

This new accounting standard will be effective from 2020. The IFRS 9 standard provides a new set of regulations that the new loss provisioning process will need to satisfy.



Figure 1, From IAS39 to IFRS 9

In particular, the new impairment rules require that the lifetime credit risk of an account be assessed at each model run to determine if there has been a significant increase in credit risk since origination. For accounts where the credit risk has significantly increased (including defaults) the lifetime expected credit losses must be used. If the credit risk has not significantly increased, then only credit losses resulting from expected defaults in the next 12 months must be used. The approach is outlined in terms of a stage classification accounting for significant increase in credit risk as a pivotal element of IFRS 9.

**Table 1: Stages under IFRS 9**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Stage 1** | **Stage 2** | **Stage 3** |
| **Stage description** | Includes accounts for whom no significant increase in credit risk since initial recognition has been observed | Includes accounts whose credit risk has significantly increased since initial recognition but no objective evidence of impairment has been observed, with a rebuttal presumption that this occurs when the account reaches 30 days past due | Includes accounts where the objective evidence of impairment has been observed |
| 12-month expected credit losses, i.e. credit losses due to default events within subsequent 12 months, are recognised in balance sheet | Lifetime expected credit losses (LTECL) i.e. credit losses due to default events spanning the (expected) lifetime of the facility, are recognised in balance sheet | Lifetime expected credit losses (LTECL) are recognised in balance sheet |
| **Expected loss** | IFRS 9 guidelines require to assess the 1 year expected credit losses without prescribing the tangible estimation procedure    , , , represent marginal PD, EAD and LGD at time *t* | IFRS 9 guidelines require the lifetime expected credit losses without prescribing the tangible estimation procedure   Where, , , , represent marginal PD, EAD and LGD at time *t* and *T* represents the remaining lifetime of the account | Expected loss of a defaulted client given the loss rate, i.e. the shortfall in net present value of expected cash flows versus the carrying amount of the loan |

Table 1, Stages under IFRS 9

The exemplary expected loss (EL) assessment in

Table 1 listed above is based on an estimate of 1 year parameters probability of default (PD), loss given default (LGD) and exposure at default (EAD) for stage 1 and multi-year PD, LGD and EAD assessments for stage 2 including a discount factor to the reporting date.

# Definition

| **Term** | **Acronym** | **Description** |
| --- | --- | --- |
| Days Past Due | DPD | The number of days that an account is currently in arrears |
| Delinquent |  | An asset is described as delinquent if it is associated with any amount of arrears |
| Expected Loss/Expected Credit Loss | EL/ECL | Interchangeable terms. EL = PD\*EAD\*LGD |
| Exposure at Default | EAD | Exposure at Default (EAD) is defined as the expected amount drawn by borrowers at the time of default. |
| Probability of Default | PD | Probability of default (PD) is the risk that the borrower will be unable or unwilling to repay its debt in full or on time. The risk of default is derived by analyzing the obligor’s capacity to repay the debt in accordance with contractual terms. PD is generally associated with financial characteristics such as inadequate cash flow to service debt, declining revenues or operating margins, high leverage or declining liquidity |
| Default customer |  | Default customer is customer who failed to make on-time repayment (>= minimum payment rate) of their loans for more than ninety consecutive days or three months. |
| Conditional Prepayment Rate | CPR | Conditional Prepayment Rate (CPR) is the annualized percentage of the mortgage expected to prepay in each period. For example, if CPR is 5%, it means that 5% of mortgage is expected to prepay within the period. The focused population is the group of opening accounts at the end of time frame. |
| Lifetime Expected Credit Loss | LTECL / LEL | The Expected Credit Loss over the behavioural lifetime of an asset |
| Beta distribution |  | The beta distribution is a family of continuous [probability distributions](https://en.wikipedia.org/wiki/Probability_distribution" \o "Probability distribution) defined on the interval [0, 1] [parametrized](https://en.wikipedia.org/wiki/Statistical_parameter" \o "Statistical parameter) by two positive [shape parameters](https://en.wikipedia.org/wiki/Shape_parameter" \o "Shape parameter), denoted by *α* and *β*, |

Table 2, Term definitions and descriptions

# Probability of Default

Probability of default (PD) is the risk that the borrower will be unable or unwilling to repay its debt in full or on time. The risk of default is derived by analyzing the obligor’s capacity to repay the debt in accordance with contractual terms. PD is generally associated with financial characteristics such as inadequate cash flow to service debt, declining revenues or operating margins, high leverage or declining liquidity.

With the new IFRS9 loan loss provision, loans are classified in three stages: stage 1 – initial recognition (yet to be impaired), stage 2 – significant increase in credit risk, and stage 3 – objective indicators of impairment. For loans in stage 1, banks need to estimate 1-year expected credit losses. On the other hand for loans in stage 2 and 3, banks need to provide provision and thus estimate expected credit losses for the whole lifetime of the loans. Both the 1-year and lifetime expected credit losses estimation shall reflect the banks’ forward looking macro-economic view.

KBank’s PD estimation for IFRS9 is modelled according to the following principles: (i) the PD estimation for IFRS9 should be point-in-time (PIT) and reflect current market conditions, (ii) the PD estimates should use structural and behavior information, and (iii) estimation of PD should include the forward looking aspect of the macroeconomic outlook specific to particular sector. For loans in stage 1 and 2, the modelled probability of default will be over 12 months and lifetime respectively. For stage 3, the probability default will be at 100%.

KBank leverages existing behavioral scorecard to construct a new credit rating system (supermaster rating) and then use it to create the probability of default term-structure of each obligor up to its maturity based on a continuous time homogeneous or non-homogeneous Markov transition matrix and then incorporate systematic risk into PD term-structure model via Vasicek model for calculating joint loss distribution of bank exposures (Vasicek, 2002).

This section outlines step by step Probability of Default (PD) model development methodology for the champion model and states its compliance with the current IFRS9 regulatory requirement. It also documented all the technical difficulties that the model developer experienced undergoing these model development steps.

## Scope

This probability of default model should be used to create the PD term structure for all valid stage-1 and stage-2 housing loan instruments. For stage-3 instruments, they are automatically assigned to be at 100%.

## Methodology Review

KBank leverages existing behavioral scorecard to construct a new credit rating system (supermaster rating) and then use it to create the probability of default term-structure of each obligor up to its maturity based on a continuous time homogeneous or non-homogeneous Markov transition matrix and then incorporate systematic risk into PD term-structure model via Vasicek model for calculating joint loss distribution of bank exposures (Vasicek, 2002).

The existing behavioral scorecard is regarded by the bank as a feeder model to the TFRS9 probability of default model. Hence, this document does not extend to the model development and validation of the current behavioral scorecard.

For KBank retail portfolio, we explore a way to construct PD term structure by applying Markov state transition model. The Markov property represents the assumption that the evolution of credit migration is independent of the past credit migration history. In other words, the probability of migration from rating class i to rating class j does not depend on any information from the past, only the current rating matters. To find out the best credit migration for our corporate lending, we consider several alternative approaches: (i) discrete time Markov transition, (ii) homogeneous continuous time Markov transition, and (iii) nonhomogeneous continuous time Markov transition.

### TTC PD: Discrete Time Markov Transition (DTMT)

We first consider a discrete time Markov chain. Let be a discrete-time transition matrix

where denotes the probability of change in rating at the beginning of a year to rating class at year’s end. In total, we have ratings (states). The final rating represents a default state.

A useful property of the time homogeneous Markov transition matrix is the fact that the *n*-year transition matrix is simply given by the *n*th power of the one-year transition matrix , denoted by i.e.

[Matrix multiplication times]

The cumulative nth year default probabilities for a starting rating class is the last column of ,

The model assumes that probability of state (rating) transition only depends on the current information (current rating) i.e. rating transition exhibit Markov property. In reality, the rating migration needs not to be stationary or time-homogeneous and the maturity of loans need not be a yearly increment. We explore further to identify a more suitable approach to model the probability of default term structure.

### TTC PD: Homogeneous Continuous Time Markov Transition

A continuous time, time-homogeneous Markov chain is predicated in terms of a symmetric generator matrix (for possible transitory states)

where for all , and for

The transition matrix for time interval is given by

where denotes the matrix exponential

However, there is evidence that the assumption of time-homogeneity is very likely not precisely true for real credit ratings migration. We therefore consider nonhomogeneous time Markov chain and would like to find that best represent the dynamics of PD migration. That is, we no longer assume that the transition rates  are constant over time. This will be done using parametric approach.

### TTC PD: Non Homogeneous Continuous Time Markov Transition

We replace the homogeneous generator for the time interval by the time-dependent-generator:

where X denote matrix multiplication and is the diagonal matrix in with elements of as follow

The function is defined as follow

where ,are non-negative constant and time .

Finally, we decide to explore the possibility of applying a nonhomogeneous continuous time Markov transition to forecast the dynamics of PD migration for our lending portfolio, because it better reflects seasoning effects and improves the fit, especially on the lowest and highest ratings.

### PIT PD: Vasicek (2002) and Credit Adjustment Factor

Previously, we derive TTC PD term structure or accumulative PD for, say, n years. The forward PD is then calculated and defined as the probability of the loan has defaulted by the end of year n, minus the probability of the loan had already defaulted before the beginning of year n. That is, the forward PD is the difference between the n-year and the n-1 year accumulative PD.

Our approach to convert forward PD to PIT PD is through the application of the Vasicek (2002) equation using a Cycle Adjustment Factor (CAF). The PIT PD for year t is given by:

Where is the cumulative standard normal distribution function

is the cumulative inverse standard normal distribution function

is the asset-segment correlation

or Cycle Adjustment Factor is the normalized risk factor that represents the point in the economic cycle

is the forward PD for an individual obligor/rating

is the PIT PD for an individual obligor/rating

We next find the fitted CAF model by seeking relationship between credit index (Z) and macroeconomic factors. That is, we regress the normalized historical credit index against the historical macroeconomic factors. The macroeconomic factors are assumed to be exogenous to the model and are listed in

Table 3 together with their expected signs.

|  |  |  |
| --- | --- | --- |
| No. | Macroeconomic Factors | Expected Sign |
| 1 | Private Consumption | - |
| 2 | Investment | - |
| 3 | House Price Index | - |
| 4 | Unemployment Rate | + |
| 5 | Diesel Price | - |
| 6 | Household Debt | + |

Table 3, List of macroeconomic factors and their expected signs

*Grouping criteria and asset correlation*

Credit Adjustment Factors (CAF) in different homogeneous asset classes are naturally different based upon their economic drivers. Since commercial portfolio is based on systematic factors in industry, we develop CAF by Occupation: Salary Earner and Entrepreneur.

To develop the econometric model, a direct regression of the observed actual default against the macroeconomic factors is performed.

The estimated regression equation is

Where measure the “credit cycle” at time t and is the OLS estimates which minimize the sum of squared error. In addition, the macroeconomic factors that are selected into the final fitted CAF model should fulfill the following conditions:

1) Its coefficients should be significant (p-value < 0.05)

2) Its coefficients should be consistent with economic theory as well as improve the overall model’s performance (e.g. adjusted R2).

Finally, given the forecasted macroeconomic factors, we can calculate CAF for the forecasted period. In conjunction with forward PD estimates, we then calculate forward-looking PIT PD using the above-mentioned Vasicek equation where is the best fit between actual and predicted model.

### TTC & PIT PD: Survival Analysis

KBank explore another fitting methodology the “Survival Analysis” to model the probability of default term structure. We attempt to rectify away from the Markov property assumption which is doubtly valid especially for mortgages instrument where the probability default and rating transitions should be highly sensitive to past behavior.

Within strands of survival analysis, we specifically consider the Cox Proportional Hazard model. The Cox PH model evaluates simultaneously the effect of several risk factors on survival. The model itself is semi-parametric, robust, and applicable with variety of data situation with a capability to incorporate censored data. Within the context of credit risk, the Cox PH model is expressed by the hazard function

Where – hazard rate of a default event happening at a particular time

– Predictor variables or covariates

– Coefficients that measure the impact of covariates

– Baseline hazard rate

We employ a slightly enhanced version of the above model i.e. the extended Cox model as a challenger model to estimate the hazard rate of a mortgage instrument. The extended Cox model is expressed by the following hazard function

Where is the baseline hazard rate for each strata

are static and are dynamic covariates

The relationship between hazard rate from extended Cox model and survival probability (curve) are defined as follow

Where is the baseline survival rate for each strata

## Model Development Approach

For PD term structure model, KBank believe that the Markov state transition model family is the appropriate methodology for the construction of PD term structure. The inputs of the model and the development exercise align with the bank’s current data structure. The output of model also aligns with current business model and TFRS9 risk requirements in term of generating the default term structure.

KBank explore the possibility of instigating the non-homogeneous continuous time Markov state transition model (NHCTMT) as a champion model for Housing Loan PD term structure. But depending on the results of the curve adjustment optimization, KBank may opt to revert to the homogeneous model instead. Additionally, KBank also attempt to use the survival analysis methodology to create the default term structure as a challenger model.

Figure 1, Steps of PD term structure model development

This section briefly outlines the steps that have been taken in the development of the champion model (NHCTMT) for the TFRS9 probability of default term structure.

The model development process starts with data gathering and sample preparation. We used a long history of default performance, B-score and TDR-score in the derivation of our supermaster rating and 1-year through-the-cycle probability of default. This will be elucidated in detail in the next section of this document.

After the supermaster rating and scale has been well developed, we re-apply the rating criteria to our historical observations and observed the 1-year rating migrations over different observation periods. We constructed our through-the-cycle 1-year state transition probability based on the average migration rates across years of observation. The next step involves human intervention in order to smooth out and ensure that the TTC rating migration matrix comply with the necessary conditions of the NHCTMT model (will be discussed later).

After the TTC migration matrix has been finalized, we transform it to the generator matrix. We use this generator matrix together with an optimization algorithm to find appropriate curve adjustment parameters and . The objective of the optimization algorithm is to find the curve adjustment parameters such that the PD curves are the closest to and still cover the historical default rates across all ratings.

## Data Management

### Overview of Input Data Set

The first step of our model development is the derivation of super master scale. We gathered Housing Loan portfolio data and B-score of each instrument from 2012-01 to 2016-05.

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Field\_Name | Existing/New | Description |
| 1. | POS\_DT | Existing | Data as of observation point |
| 2. | LPM\_NO | Existing | Customer ID |
| 3. | ASSET\_CLASS | Existing | Asset class of customer is to identify Good/Bad customer. If the asset class contains “NPL” then it is defined as bad if not then it is defined as good. |
| 4. | B\_score | Existing | Customer Level B-Score (minimum from account level) |
| 5. | Supermaster\_rating | New | Supermaster rating derived from 2.1 |

Table 4, Input data

From the data we observe portfolio actual default and construct a new super master rating. Subsequent to the completion of the supermaster rating, we observe a rating migration of the same observation and performance period. This process is outlined in the figure below.



Figure 2 Observation and outcome of default event and rating transition

We also utilize occupation data to explore the possibility of segmentation of the migration matrices.

### Sample Design

The observation point is every quarter starting from January 2012 to May 2016. The performance period for observation of actual default or rating migration is 12 month from the observation point i.e. January 2012 to May 2017.

### Data Cleansing and Exception Handling

In the construction of our development sample, at every observation point we exclude records using the following rules

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Issue** | **Detail** | **Handling** |
| 1 | NPL record | The record/customer is NPL at the observation | Remove from sample |
| 2 | Open less than 3 months | The record/customer is new and there is no behavioral score | Remove from sample |
| 3 | No performance | The record/customer with no performance i.e. closure, write-off etc. | Remove from sample |

Table 5, Data cleansing and exception handling rules

### Final Development Sample

This section show step by step derivation of our final model development sample. Starting from the entire data set, working the way through record exclusion and thus the final development sample and associated bad rate.

Tables below show the number of customers in each of the observation point before and after exclusion.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Asset Class/Month | Jan-12 | April-12 | July-12 | Oct-12 | Jan-13 | April-13 | July-13 | Oct-13 |
| Total | 48,888 | 51,475 | 55,609 | 59,003 | 60,216 | 62,944 | 66,521 | 69,249 |
| Exclusion | | | | | | | | |
| NPL | 738 | 772 | 896 | 965 | 1,114 | 1,143 | 1,278 | 1,379 |
| MOB<3 | 1,737 | 3,122 | 5,243 | 5,130 | 2,498 | 3,358 | 4,882 | 4,007 |
| No Performance | 637 | 676 | 696 | 729 | 706 | 772 | 846 | 969 |
| Final Model Development Sample | | | | | | | | |
| Total | 45,776 | 46,905 | 48,774 | 52,179 | 55,898 | 57,671 | 59,515 | 62,894 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Asset Class/ | Jan-14 | April-14 | July-14 | Oct-14 | Jan-15 | April-15 | | July-15 | | Oct-15 | | Jan-16 | | April-16 | |
| Total | 69,959 | 71,851 | 74,062 | 77,154 | 79,742 | 83,041 | | 86,622 | | 89,389 | | 91,980 | | 95,200 | |
| Exclusion | | | | | | | | | | | | | | | |
| NPL | 1,626 | 1,484 | 1,443 | 1,655 | 1,767 | 1,762 | | 2,185 | | 2,751 | | 3,487 | | 3,759 | |
| MOB < 3 | 1,824 | 2,560 | 3,230 | 4,303 | 4,103 | 4,304 | | 5,261 | | 4,512 | | 3,994 | | 4,351 | |
| No Performance | 869 | 873 | 947 | 1,017 | 1,077 | 1,231 | |  | |  | |  | |  | |
| Final Model Development Sample | | | | | | | | | | | | | | | |
| Total | 65,640 | 66,934 | 68,442 | 70,179 | 72,795 | 75,744 | | 79,176 | | 82,126 | | 84,499 | | 87,090 | |



Table 6, Observations and development samples

The performance of the development sample (i.e. number of bad customer within the defined performance period) is as per the table below. Please note that there are two default performance definitions for Housing Loan portfolio. The first definition is 90+ DPD as reflected by BOT class. The second definition is TDR loans behavioral default. For detail on the definition, please refer to the definition section of this document.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Asset Class | % Bad | | | | | | | |
| Q1 Jan-12 | Q2  Apr-12 | Q3 July-12 | Q4  Oct-12 | Q1 Jan-13 | Q2  Apr-13 | Q3 July-13 | Q4  Oct-13 |
| All | 1.16% | 1.16% | 1.26% | 1.28% | 1.52% | 1.10% | 1.08% | 1.30% |
| PL Normal |  |  |  |  |  |  |  |  |
| SMA/SMQ |  |  |  |  |  |  |  |  |
| TDR |  |  |  |  |  |  |  |  |
| Watchlist/Reschedule |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Asset Class | % Bad | | | | | | | | | |
| Q1 Jan-14 | Q2  Apr-14 | Q3 July-14 | Q4  Oct-14 | Q1 Jan-15 | Q2  Apr-15 | Q3 July-15 | Q4  Oct-15 | Q1  Jan-16 | Q2  Apr-16 |
| All | 1.26% | 1.31% | 1.87% | 2.31% | 4.13% | 5.01% |  |  |  |  |
| PL Normal |  |  |  |  |  |  |  |  |  |  |
| SMA/SMQ |  |  |  |  |  |  |  |  |  |  |
| TDR |  |  |  |  |  |  |  |  |  |  |
| Watchlist/Reschedule |  |  |  |  |  |  |  |  |  |  |



















Table 7, Default performance of the development sample

## Model Development

### Super master Scale and Rating

The first step in our PD term structure model development is the construction of supermaster rating and supermaster scale PD. For current loan and TDR customer, we use scoring model to create ratings and thus assign a suitable rating for each customer. For other asset classes, we directly observe the long run historical default rate within each asset class and assign them to an appropriate rating in the supermaster scale.



Figure 3, Overview of Data Input

The outcome of this exercise is the supermaster rating and scale as shown below. The PD in each rating are calibrated to the long run default performance up to June 2016.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Super master Rating** | **Asset flag** | **B-Score/TDR-Score Range** | **PD Master Scale** | **201612** |
| **30/1/2018** | **Actual DR** |
| 1 | PL Normal | >654 | 0.25% | 0.31% |
| 2 | PL Normal | 571-654 | 0.76% | 0.84% |
| 3 | PL Normal | 471-570 | 1.40% | 1.52% |
| PL Normal | (NEW) |
| 4 | PL Normal | 395-470 | 4.00% | 4.46% |
| 5 | PL Normal | 311-394 | 7.26% | 7.19% |
| Reschedule | . |
| 6 | PL Normal | 221-310 | 11.15% | 9.98% |
| Watch list | . |
| TDR | >590 |
| 7 | PL Normal | 41-220 | 25.89% | 23.94% |
| TDR | 561-589 |
| TDR | 521-560 |
| 8 | TDR | 478-520 | 49.91% | 34.01% |
| PL Normal | <41 |
| 9 | SMA/SMQ | . | 53.04% | 47.64% |
| TDR | <478 |

Table 8,Super master Scale for PD

### TTC Model

After the super master rating and scale are finalized. We re-apply the rating criteria to our sample and observe the migration of supermaster rating. Within our sample, we observed a weighty dissimilarity of transition rates and default trends for customer with different occupation i.e. salary earner and entrepreneur. Furthermore, the Bank’s acquisition policy framework for these two groups are vastly different. The definition of occupation is grouped as follow: entrepreneur and freelance are flagged as “Entrepreneur”, otherwise they are “Salary Earner”.

By taking a simple average across each observation point, we obtain the following obtain normalized long-run average (observed) transition matrix or through-the-cycle (TTC) transition matrix.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Super master Rating | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Default |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| Default |  |  |  |  |  |  |  |  |  |  |

Table 9,Observed transition matrix for Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Super master Rating | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Default |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| Default |  |  |  |  |  |  |  |  |  |  |

Table 10, Observed transition matrix for Entrepreneur

It is necessary that our transition matrices must possess the property of diagonal dominance. The probability of rating transition from the initial state should be a bell curve i.e. the diagonal transition probability (identical initial and final state) should be the highest transition rate and then subsequently decreasing according to the distance from the starting state (rating). We smooth out our transition rates to obtain such property using to the methodology from [4] with aiming to replicate the external rating properties from Moody’s and S&P’s.

The key concept of transition smoothing is to maintain the average movement from the observed matrix while lowering a variance in each rating. In addition, the left most or the right most rating should be the highest average movement since they can be only upgraded or downgraded. On the other hand, the middle rating should be the lowest average movement since there is a same probability to move upward or downward. Having said that, the adjustment process involves of expert judgment and is a manual process.

Average movement for rating can be calculated through

Standard deviation for rating is then expressed as follow

It is then possible to compare the average movement and standard deviation of transition of the observed and the adjusted transition matrix‘s average movement. Note that the goal of the smoothing process is to find a transition matrix with good property while minimizing the changes to the average movement and standard deviation from the original matrix. The outcome of this stage are the adjusted and normalized transition matrices (final) as follow

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 11, Smoothed transition matrix for Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 12, Smoothed transition matrix for Entrepreneur

The new matrices have the following average movement and standard deviation which are comparable to the original observed matrices.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rating | Average Movement | | | Standard Deviation | | |
| Observed | Smooth | Smooth  (Normalized) | Observed | Smooth | Smooth  (Normalized) |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |

Table 13, Pre and post adjustment: average movement and S.D. Salary Earner

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rating | Average Movement | | | Standard Deviation | | |
| Observed | Smooth | Smooth  (Normalized) | Observed | Smooth | Smooth  (Normalized) |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |

Table 14, Pre and post adjustment: average movement and S.D. Entrepreneur

### Generator matrix

This part will describe how to derive the generator matrix as continuous time transition matrix from the matrix M (the discrete matrix) within the acceptable range of error. Please see the methodology review section on the property of the generator matrix.

We use a library in R called “ctmcd” and “expm” in order to transform matrix M to matrix Q. More specifically, we choose “gmDA” function available in “ctmcd” library to find generator matrix Q and “expm” is the function to calculate Taylor series of matrix exponential. In addition, we need to find the generator matrix Q such that it minimizes the error between the discrete matrix and the continuous time matrix (after exponential or natural logarithm)

Where is our acceptable level of error

With the default “gmDA” function, we obtain the following value of the objective function: 0.00285 which is intolerable. In order to reduce this error, we use a scaling matrix (a diagonal matrix) to adjust the generators slightly.

The new objective function becomes

The optimization yields the following value of and the desired generator matrices for salary earner and entrepreneur respectively

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 15, Adjustment Marix A, Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 𝐴 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 16, Adjustment Marix A, Entrepreneur

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 17, Final Generator matrix for Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| D |  |  |  |  |  |  |  |  |  |  |

Table 18, Final Generator matrix for Entrepreneur

### Non-homogeneous curve adjustment

In this section, we will explain how to obtain an optimal choice of alpha and beta to best fit the model. Suppose that is output of taking an exponential of the non-homogenous transition matrix as follow

The output is highly dependent on the curve adjustment function as reviewed in the methodology review section. Mathematically speaking, one can construct the optimization equation to find the scalar of alpha and beta such that the modelled probability term structure is the closest to the actual cumulative default rate at any time i.e.

Where, is the rating in consideration

is the historically cumulative default rate of each rating at time

is the PD for each rating from the generator matrix Q at time .

The inequality constraint is to ascertain that probability of default from the model will always be higher than the actual DR at any point on the term structure. We adopt R programming with function “solnp” available within the “RSolnp” library as our optimization tool.

The  is obtained from observing an average 5 year historically marginal default rate of performing account in the sample between 201201-201206 and then sum up each year to obtain cumulative. The default definition is the same as in the previous exercise i.e. 90+ DPD and behavioral default (from 2016 onwards). For example, for observe point 201201, the marginal default rate for rating 1 can be illustrated as per the following figure



Figure 4, Observable default rate term structure

The cumulative default rate for each rating for salary earner and entrepreneur that will be use in the calibration of our PD term structure is as follow

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Supermaster Rating | Year | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |

Table 19, Default term structure for Salary Earner

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Supermaster Rating | Year | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |

Table 20, Default term structure for Entrepreneur

By solving optimization above, we did not obtain satisfactory term structure curve. There was a big gap on the first year especially on the lower rating between PD model and actual default. Furthermore, we were unable to force the term structure to sit above the actual default for higher rating. In this quest, we revert back to using the homogenous version of the continuous time markov chain which is a convertism choice. We still use the following curve adjustment parameter (equivalent to a homogenous chain)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Curve Adjustment Parameter | Supermaster Rating | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

Table 21, Curve adjustment parameter for Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Curve Adjustment Parameter | Supermaster Rating | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | D |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

Table 22, Curve adjustment parameter for Entrepreneur

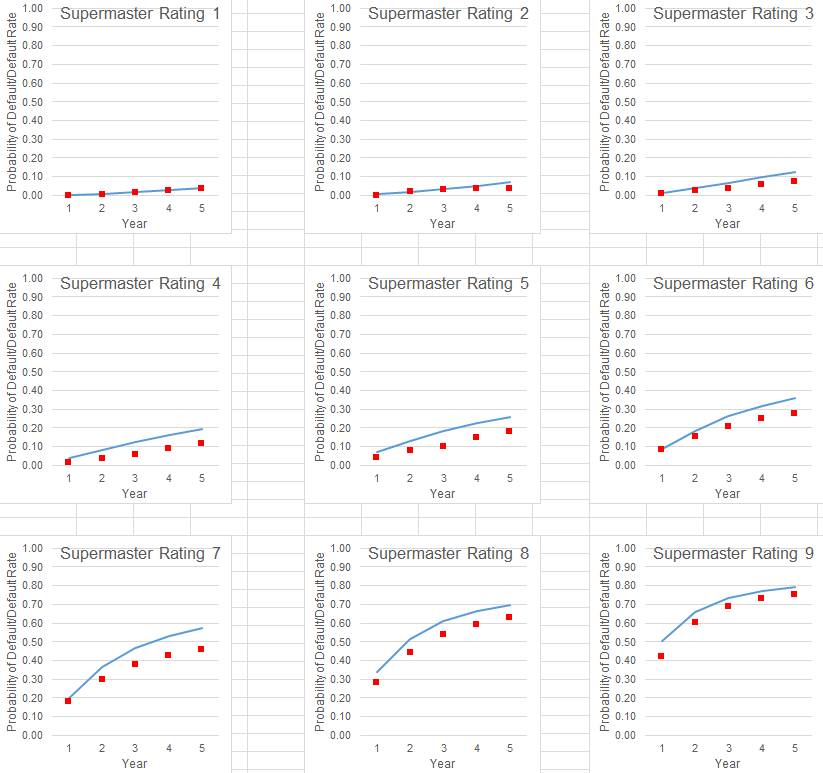
Subsequently, one can obtain the cumulative TTC PD term structure for 5 years from the model follow

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Supermaster Rating | Actual Cumulative Default Rate | | | | | | Model TTC Cumulative Default Rate | | | | | |
| Year | | | | | | Year | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 4 | 5 |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |  |  |

Table 23, Cumulative default rate: actual vs. model, Salary Earner

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Supermaster Rating | Actual Cumulative Default Rate | | | | | | Model TTC Cumulative Default Rate | | | | | |
| Year | | | | | | Year | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 4 | 5 |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |  |  |

Table 22, Cumulative default rate: actual vs. model, Entrepreneur



### PIT Model

In this section, we will explain how to convert TTC PD obtained from the previous steps to PIT PD via the Vasicek equation. There are 2 components namely the and that would need to be identified. We first construct CAF based on portfolio segmentation which is by occupation namely salary earner and entrepreneur which correspond to the segmentation done in the previous steps.

The data preparation is to identify which macro variable is suited for each sub-segmentation and it can be generated by seeking the relationship between actual default and macro-economic index via R programming. Specifically, the observation period is quarterly from Jan-2012 to April-2016 and the performance period is 1 year sliding window from Jan-2013 to April-2017. The good and bad definition is exactly the same as curve fitting in Non-homogenous section.

The historical macro variable index is based on the Bank’s internal organization research or “K-Research”. Since PIT PD is a forward-looking concept then the point of macro-economic index will be the next 12 months from observation point. The greater detail is shown as the following table:

|  |  |  |
| --- | --- | --- |
| No. | Name | Description |
| 1. | Date | Data monthly as of observation point from 200901 to 201603 |
| 2. | DR\_Salary\_Earner | 1 year actual default rate of Salary Earner from Date |
| 3. | DR\_Self\_Employed | 1 year actual default rate of Entrepreneur from Date |
| 4. | GDP\_2002 | Historical Thai GDP data as of (Date + 12 months) |
| 5. | PRIVATE\_CONSUM | Private Consumption index data as of (Date + 12 months) |
| 6. | INVESTMENT | Investment index data as of (Date + 12 months) |
| 7. | GOV\_SPEND | Government spending index data as of (Date + 12 months) |
| 8. | EXPORT | Export index data as of (Date + 12 months) |
| 9. | IMPORT | Import index data as of (Date + 12 months) |
| 10. | UNEMPLOY\_RATE | Unemployment rate index data as of (Date + 12 months) |
| 11. | HEADLINE\_CPI | Headline CPI index data as of (Date + 12 months) |
| 12. | CORE\_CPI | Core CPI index data as of (Date + 12 months) |
| 13. | DIESEL\_PRICE | Diesel Price index data as of (Date + 12 months) |
| 14. | MLR | MLR index data as of (Date + 12 months) |
| 15. | HOUSE\_PRICE | Houseprice index data as of (Date + 12 months) |
| 16. | HOUSEHOLD\_DEBT | Household debt index data as of (Date + 12 months) |
| 17. | CCI | CCI index data as of (Date + 12 months) |

Table 24, Input data for macro-overlay model

We also randomly spit sample into testing and training i.e. 80 to 20 respectively. After that, we fit a linear model to the time series data to establish the relationship between macro variables and the default rates over time. The selections of independent variables are done via a backward selection process. However, the final set of macro-economic variables selected into the final model is based on economic intuition, predictive significance of the independent variable and a tolerable level of model accuracy (Adjusted Rsquare).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Product | Occupation | *β0* | *βUnemp.Rate* | *βHHDebt* |  |  | *Adjusted* |
| Housing Loan | Salary Earner |  |  |  |  |  | 30-60% |
| Entrepreneur |  |  |  |  |  | 30-60% |

Table 25, Final coefficient of macro-overlay model

After obtaining CAF, the next part is to calculate that is the best fit to the model. In other word, we put the actual default as PD PIT and PD TTC as forward-looking PD with calculated CAF from above then one can solve optimization problem to obtain . However, due to data availability with more prudential correlation, we use a benchmark from Moody’s for salary earner and empirical observation for entrepreneur result of 4 and 10 percent respectively. This is because we believe that entrepreneur is more sensitive to systematic risk than salary earner.

|  |  |  |
| --- | --- | --- |
| Product | Occupation | Correlation |
| Housing Loan | Salary Earner | 4% |
| Entrepreneur | 10% |

Table 26, Correlation factor

## Pre-Validation

### TTC Model

For the 1-year TTC PD in the master scale we have validated by using a binomial test. The test results are as follow

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Supermaster Rating** | **Asset\_flag** | **B-Score/TDR-Score Range** | **PD MasterScale 30/1/2018** | **Critical PD30/1/2018** | **Actual DR 201612** | **Binomial Test** |
| 1 | PL Normal | >654 | 0.25% | 0.31% | 0.31% | Accept |
| 2 | PL Normal | 571-654 | 0.76% | 0.84% | 0.84% | Accept |
| 3 | PL Normal | 471-570 | 1.40% | 1.57% | 1.52% | Accept |
| PL Normal | (NEW) | Accept |
| 4 | PL Normal | 395-470 | 4.00% | 4.47% | 4.46% | Accept |
| 5 | PL Normal | 311-394 | 7.26% | 8.17% | 7.19% | Accept |
| Reschedule | . | Accept |
| 6 | PL Normal | 221-310 | 11.15% | 12.84% | 9.98% | Accept |
| Watch list | . | Accept |
| TDR | >590 | Accept |
| 7 | PL Normal | 41-220 | 25.89% | 28.07% | 23.94% | Accept |
| TDR | 561-589 | Accept |
| TDR | 521-560 | Accept |
| 8 | TDR | 478-520 | 49.91% | 51.64% | 34.01% | Accept |
| PL Normal | <41 | Accept |
| 9 | SMA/SMQ | . | 53.04% | 55.21% | 47.64% | Accept |
| TDR | <478 | Accept |

Table 27, Binomial test of Super master Scale

For the TTC PD term structure, the generator matrix and the curve adjustment parameters are calibrated to the PD term structure from Jan-2013 to Jan-2017. At the time of development, we have yet to identify an appropriate measure to validate the performance of the PD term structure. Specifically for housing loan portfolio, we ensure that the PD curves are conservative i.e. higher than the actuals for all ratings and at every point in time.

|  |  |
| --- | --- |
| Curve Adjustment | SSE |
| Salary Earner | <0.2 |
| Entrepreneur | <0.2 |

|  |  |
| --- | --- |
| Curve Adjustment | MAPE |
| Salary Earner | <100% |
| Entrepreneur | <100% |

### PIT Model

The performance of the (linear model) is validated via adjusted. Both the salary earner and entrepreneur has an acceptable range of adjusted as follow

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Product | Occupation | β0 | βUnemp.Rate | βHHDebt.Lag3M | βGDP | Adjusted |
| Housing Loan | Salary Earner |  |  |  |  | 30-60% |
| Self employed |  |  |  |  | 30-60% |

Table 28, Adjusted R^2 for macro-overlay models

# Exposure at Default

EAD is one of the major components which are used to calculate credit risk capital and provision. Based on nature of the products, we can discriminate them into 2 product types which are Term Loan and Revolving Loan. For Retail Credit products, term loan is consisting of Housing Loan, Staff Loan, Other Secured Loan and K-Leasing. EAD-TFRS9 model will be segmented based on Term Loan and Revolving Loan as shown in Figure1 below. The first one is model for loan products and the other one is model for revolving products.

**Retail Product**

**B. Revolving Loan**

1. **Term Loan**

**Credit Card**

**POD**

**Housing Loan**

Other Secured Loan

**Consumer Loan**

**K-Express Cash**

Other Unsecured Loan

**Staff Loan**

New Car

Used Car

K- Car

**K-Leasing**

Figure 5, Product Structure of Retail Product

This development document will solely concentrate on the model for term loan as housing loan is considered as a term loan product. EAD model developed in this document is used in housing loan and staff loan portfolio because both are home mortgage loan.

***A: Term Loan***

The mortgage industry of Thailand pays a major role in consumer lending business. The market is fiercely competitive. Both government and commercial banks attempt to attract borrowers by offering special interest rate programs. Several programs are created such as fixed 1-year rate, fixed 2-year rate, fixed 3-year rate, and etc. Borrowers can match the rate with their preference. Moreover, banks in Thailand allow borrowers to pay off the loan before the contractual maturity (pre-payment) without any penalty. So, pre-payment and pre-settlement significantly reduces Bank’s revenue. If banks can foresee which borrowers are certainly to make full prepayment or ones are to refinance, banks can prepare the rates matching borrowers’ behavior in terms of interest rate scheme and costs. Another advantage of predicting prepayment is that banks can more effectively allocate their provision matching borrowers’ behavioral life time.

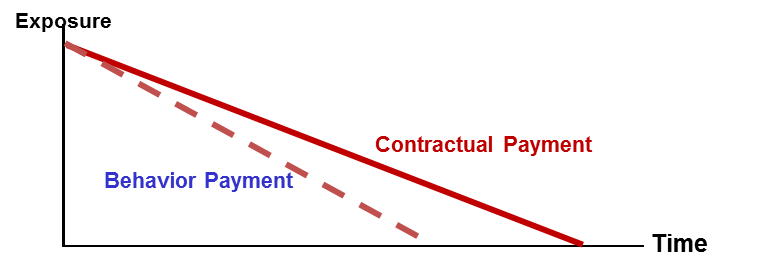
The objective of this development is to more efficiently estimate customers’ lifetime and exposure curve. In other words, banks would like to know how long the borrowers stay in the portfolio and what the exposure curve looks like.

Figure 6, Concept of Prepayment Model

As shown in Figure 4, with the effect of prepayment, the remaining of exposure would be reduced when compared to contractual exposure. The effect can be incorporated by %CPR in equation 1.

(1)

Where

= Conditional prepayment rate for month t

= The monthly contractual cash flow which generates from non-ECL part.

(2)

As described in equation 2, EL could be reduced by component. In this document, the Conditional Prepayment Rate model would be developed to assess lifetime’s expected loss.

**Conditional Prepayment Model**

The fundamental idea is to calculate Conditional Prepayment Rate (CPR) which is the annualized percentage of the mortgage expected to prepay in each period and expressed in equation 3. For example, if CPR is 5%, it means that 5% of mortgage is expected to prepay within the period. The focused population is the group of opening accounts at the end of time frame.

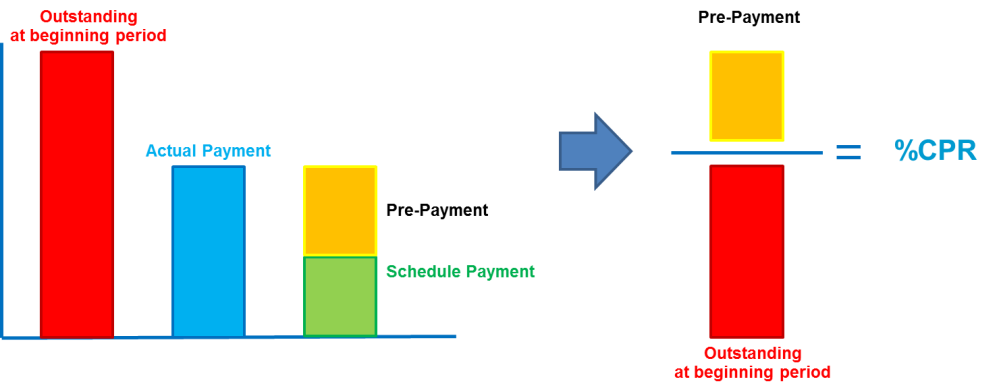


Figure 7, Conditional Prepayment Rate (CPR)

(3)

## Scope

The CPR model used to compute EAD are available to all active housing loan account except NPL customers. This process can be done by dividing the CPR model into 6 segments according to risk grade and month on book.

## Methodology Review

### Through-the-Cycle CPR by segmentation

The constant CPR in each segment is calculated based on historical average data.

### Point in time CPR

To comply with IFRS9, Kbank develop a Point in time CPR model which takes into account the macroeconomics factors that affect Kbank’s customer CPR. The appropriate macroeconomic factors will be selected and the linear regression will be performed to find the specific value of coefficient (beta) for each factor.

## Model Development Approach

This section will provide a step by step of how the EAD is calculated

CPR model development approach is done step by step as below

1. Sample Design and Data Preparation: In this step, data required for created CPR model are collected including data of all customers who hold KBank’s home loan product both demographic data, term structure, cash flow and behavioral data.
2. Data Cleansing: When modeler investigated cash flow data, some data issue are found and data cleansing is required. Details of data cleansing and exclusion will be explained in Data Cleansing and Exception Handling section
3. Calculate CPR: At this step data are combined to calculate CPR for each customer.
4. Data Exploratory Analysis: Modeler did data exploratory analysis on CPR and prospect predictor variables by visualizing relation between each predictor and CPR.
5. Model Selection: This step is done by using expert judgment to create segment based on insight from data exploratory analysis. The final model is constant CPR by segment of customer.

## Data Management

### Overview of Input Data Set

The input for actual CPR calculation includes schedule payment, actual payment, month-on-book, risk grade and day past due (DPD). In summary, the data requirement for the calculation of actual CPR calculation is shown as follow:

Figure 8, Data required for actual CPR calculation

### Sample Design

For calculating conditional prepayment rate, all active accounts in Housing Loan portfolio without any data issue will be observed. The partial pre-payment rate will be observed monthly historical behavioral payment during Jan 2010 – Apr 2016

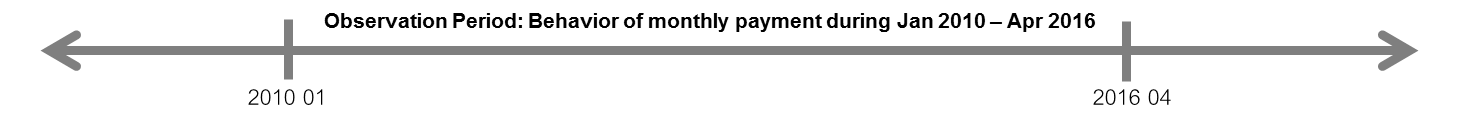


Figure 9, Development data sets periods

### Data Cleansing and Exception Handling

In order to avoid the data issue, thus, the data cleansing and exception handling process is required to filter out the undesired account.

All data issue has to be cleaned as table below;

|  |  |
| --- | --- |
| All Product HL Data | Exclusion Detail |
| Model scope exclusion | Account from RIW6100 (Status=NPL) |
| Data Issue/ Data missing | No contractual payment data (aging>0) |
| Unable to identify partial prepayment data for overdue customer |
| Outstanding Balance = 0 or Missing |
| Cash Flow = 0 or Missing |
| Payment Amount < 0 |
| Cash Flow month and date missing/error |
| Demographic information missing |

Table 29, Exclusion Rules

### Final Development Sample

From the observation period, the total remaining of account in the model was 58,353 accounts. The detail of cleansing account for each item as shown in the table below;

|  |  |  |
| --- | --- | --- |
| Item | Exclusion Detail | Accounts |
| All account HL product |  | **135,912** |
| Model scope exclusion | Account from RIW6100 (Status=NPL) | 30,541 |
| Data Issue/ Data missing | No contractual payment data (aging>0) | 5,688 |
| Unable to identify partial prepayment data for overdue customer | 12,296 |
| Outstanding Balance = 0 or Missing | 1,439 |
| Cash Flow = 0 or Missing | 18,392 |
| Payment Amount < 0 | 2,583 |
| Cash Flow month and date missing/error | 271 |
| Demographic information missing | 6,349 |
| Remaining in CPR model | | **58,353** |

Table 30, Details of cleansing account

## Model Development

### Through-the-Cycle Model

The conditional prepayment model is to create the curve of partial prepayment rate (%CPR) over a particular period which is a month on book. All variables that might affect the amount of partial prepayment are collected based on the assumption that past experience can be used to predict the future. Therefore, forecasting conditional prepayment rate (CPR) would involve several factors in many dimensions and all variables are observed based on time dimension. The list in Table 31 demonstrates all characteristics that were taken into account.

|  |  |  |
| --- | --- | --- |
| No | Variables | Category |
| 1 | Month on Book | Performance Details |
| 2 | Year on Book | Performance Details |
| 3 | Customers' Age | Personal Details |
| 4 | Gender | Personal Details |
| 5 | Highest Education | Personal Details |
| 6 | Verified Monthly Income | Income & Wealth |
| 7 | Occupation | Employment Details |
| 8 | Developer Grade | Collateral Details |
| 9 | Province of Collateral | Collateral Details |
| 10 | Loan to Value, %LTV at origination | Collateral Details |
| 11 | Behavior Risk Grade (Low Risk, High Risk) | Performance Details |
| 12 | Worst Delinquency Status in Last 3 months | Performance Details |
| 13 | Worst Delinquency Status in Last 6 months | Performance Details |
| 14 | Worst Delinquency Status in Last 9 months | Performance Details |
| 15 | Worst Delinquency Status in Last 12 months | Performance Details |
| 16 | Current Delinquency Status | Performance Details |

Table 31, List of Independent variables

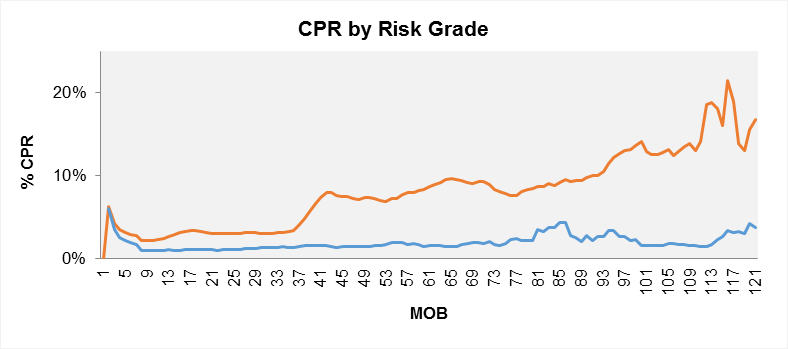
**Final Model**

To create model for TTC CPR, modeler analyzed the strength of a variable, and qualitative judgment is considered. The first chosen factor which could differentiate the rate of prepayment behavior is customer behavior risk grade. Referring to the behavior risk grade, there are 8 risk grades which are G01-G08. In this model, behavior risk grade is divided into 2 groups: Low Risk Grade and High Risk which are G01-G03 and G04-G08 respectively.

|  |  |
| --- | --- |
| Risk Level Group | Rigk Grade |
| Low risk grade | G01 - G03 |
| High risk grade | G04 - G08 |

Table 32: Risk Level Group Definition

According to historical data, the conditional prepayment rate (CPR) of low risk customer is much higher than high risk group as shown in table 32. Furthermore, %CPR of older vintage is also higher than the newer vintage. %CPR Accordingly



Low Risk Grade

High Risk Grade

Figure 10 Conditional Prepayment Rate (CPR) by Risk Grade

The final CPR model is constants for 6 segments separated by risk grade and MOB. CPR constant for each segment is weighed average CPR from the historical data of that segment. The purpose of using this model is for simplicity in strategic decision and portfolio management. CPR value for each segment is shown in Figure 11.



Figure 11 Though-The-Cycle Prepayment Rate (CPR)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Segment | DPD | Risk Grade | MOB | CPR\_TTC |
| 1 | 0 | G01-G03 | 0-36 |  |
| 2 | 0 | G01-G03 | 37-72 |  |
| 3 | 0 | G01-G03 | >72 |  |
| 4 | 0 | G04-G08 | 0-72 |  |
| 5 | 0 | G04-G08 | >72 |  |
| 6 | >0 |  |  |  |

Table 33 Though-The-Cycle Prepayment Rate (%CPR\_TTC)

### PIT Model

In order to comply with TFRS9, we have to calculate both through-the-cycle and point-in-time which use to create term structure. Macro-economic factors are incorporated in the model as shown in table 34. Using linear regression with %CPR through-the-cycle, we will regress between %CPR point-in-time as a dependent variable and macroeconomics factors as independent variables. For which we select only macroeconomics variables with expected sign. As a result, the selected macroeconomics variables with their sign, coefficient and R2 are shown in table 35.

|  |  |  |  |
| --- | --- | --- | --- |
| No | Variables | Expected Sign | Category |
| 1 | Interest Rate/ MLR | + | Macro-economic Outlook |
| 2 | House Price Index | + | Macro-economic Outlook |
| 3 | GDP | + | Macro-economic Outlook |
| 4 | Household Debt | - | Macro-economic Outlook |
| 4 | Unemployment Rate | - | Macro-economic Outlook |

Table 34, List of Independent variables

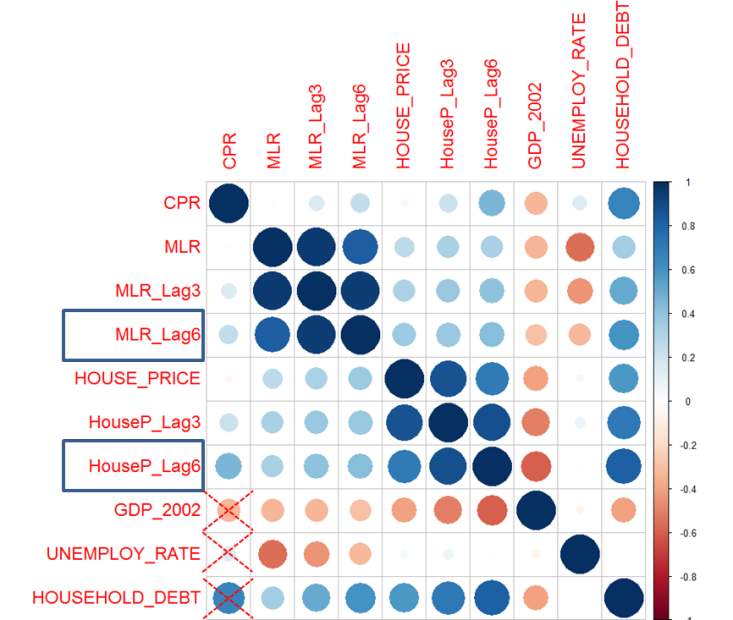


Figure 12, Correlation between %CPR and Macro-economic factors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Segment | Intercept | MLRLag6 | HousePriceLag6 | R-Square |
| 1 |  |  |  | 10-30% |
| 2 |  |  |  | 10-30% |
| 3 |  |  |  | 10-30% |
| 4 |  |  |  | 10-30% |
| 5 |  |  |  |  |
| 6 |  |  |  |  |

Table 35, CPR Model Point-in-Time

The output from partial prepayment model is incorporated into exposure at default (EAD) and becomes component of Expected Credit Loss (ECL) calculation.

## Pre-Validation

In order to validate model performance, mean squared error (MSE) will be used as the formula below.

### TTC Model

MSE Model for TTC

|  |  |  |  |
| --- | --- | --- | --- |
| **Model segment** | | | **MSE** |
| **Risk Grade** | | **MOB** |
| 1. | G01-G03 | MOB <=36 | <0.3% |
| 2. | 36< MOB <=72 | <0.3% |
| 3. | MOB >72 | <0.3% |
| 4. | G04-G08 | MOB <=72 | <0.3% |
| 5. | MOB >72 | - |

Table 36, Mean squared error for TTC model

(5)

### PIT Model

The performance of the PIT macro overlay model is as follow (R-Square)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Segment | Intercept | MLRLag6 | HousePriceLag6 | R-Square |
| 1 |  |  |  | 10-30% |
| 2 |  |  |  | 10-30% |
| 3 |  |  |  | 10-30% |
| 4 |  |  |  | 10-30% |
| 5 |  |  |  |  |
| 6 |  |  |  |  |

Table 37, R-square for PIT model

# Loss Given Default

Loss given default, LGD, can be defined as the share of a defaulted exposure that will never be recovered by the lenders. The loss given default shall be assessed in an economic sense rather than a mere accounting perspective. That said the discount effect associated with the recovery cash flow and cost associated with collecting recoveries shall be considered.

LGD amount is the total EAD amount subtracted by the expected recovery amount and plus by the collection cost. Recoveries must include both late payments and the sale price of collateral when liquidated. The discount rate used should be the cost of equity, and recoveries should be discounted back to the date of the first missed payment leading to default. Then, to get the % LGD, the LGD amount is divided by the total EAD amount as shown in the equation below,

Where,

= LGD of customer *i* at time *t* after default

= recovered cash flows and predicted cash flows of customer *i* at time *j* minus direct and indirect costs associated with the collection process

= Expected cash flows of customer *i* at time *j*

= present value function using appropriate discount rate

= outstanding of customer *i* at the time of default

Recovery cash flow is the amount of loan or obligation that will possibly be repaid to creditors in the event of a default then deducted by the direct and indirect costs associated with the collection process.

For the formula of %LGD calculation as above will be applied to both customers whose cash flow process has already finished and those with unfinished cash flow process. With unfinished cash flow process customers, expected recovery cash flow was therefore required to calculate and aggregate with the actual cash flow (see more details in section 6.5).

Then post-default state of customer will be defined which can be separated into 4 states as following:

|  |  |  |
| --- | --- | --- |
| Item | Path | Explanation |
| 1. | Self-Cured | The defaulted customer takes a certain amount of time to recover. However, No significant loss and no change in the structure or conditions of the facilities. |
| 2. | Early Cured | The defaulted customer takes a certain amount of time to recover and might not be able to fulfill his/her contractual obligations from time to time. However, No significant loss and no change in the structure or conditions of the facilities. |
| 3. | Restructuring | The defaulted customer recovers after a restructuring of his/her facilities. Usage of collateral may sometimes be part of the restructuring. Loss amount varies whereas customer relationship maintained |
| 4. | Liquidation | All facilities of the defaulted customers are liquidated, i.e. sales of loans, usage of collateral, etc. Loss amount is generally higher than that observed from restructuring. End of customer relationship. |

## Scope

Through-the-cycle LGD of home loan product is a constant, thus LGD model will be applied identically in all asset classes. For point-in-time LGD, it will be equal to through-the-cycle LGD because there is no model to adjust.

## Methodology Review

Two models were considered in the LGD for housing loan.

To develop model for LGD, modeler select 2 alternatives in consideration. The first one is to use constanct LGD for housing loan portfolio. This constant is setted by observe from historical actual LGD data. The second model in conseration as challenger model is LGD by multinomial logit overview of this model is described below.

#### Multinomial Loss Given Default

For multinomial LGD, default customers are classified into 3 ending path. The model predicts probability that each defa ult customer go into each ending path. The second step is to estimate severity of LGD for each path and for the final step, the final LGD is estimate by combining probalility with severity.

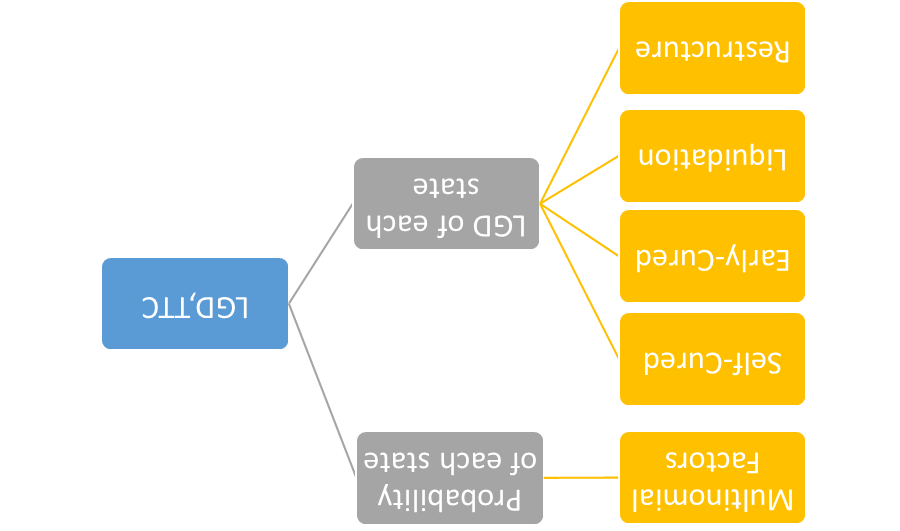


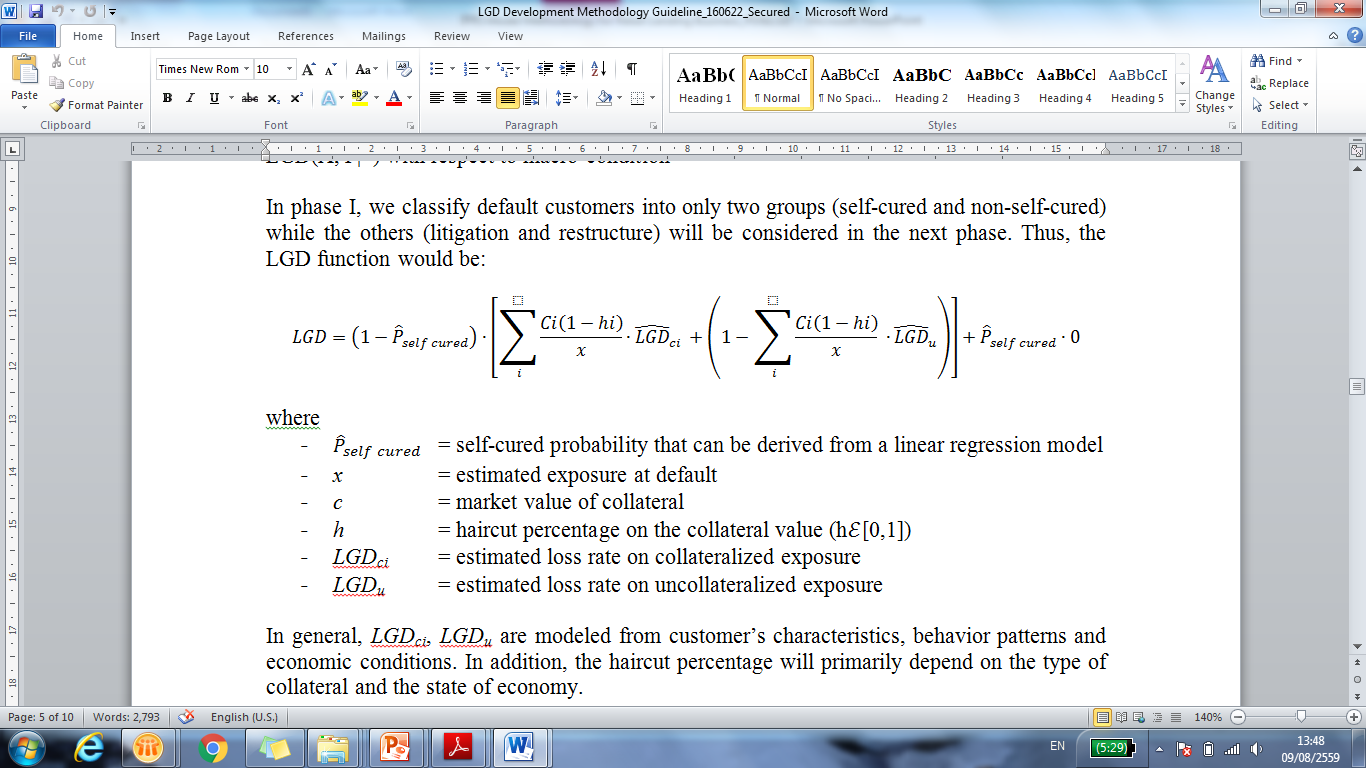
Figure 13, Concept for multinomial logit LGD

Step 1: To define probability for each state for TTC

For the first grey box, we will use multinomial factors which multinomial regression coefficient is shown in the table 3. These will use to define probability for each state from equation below.

Step 2: To define LGD for each state for TTC

Loss Severity Estimation:



The formula for calculating final LGD through-the-cycle of housing loan are presented as following

Multinomial logit LGD is considered as challenger model because data quality issues in collecting old default customers’ behaivioral data. Finally, a constant LGD is used as champion model for housing loan LGD.

## Model Development Approach

Development sample was prepared through the process of data cleansing in the first step. The group sample needed to match with relevant cash flow table source which was later on aggregated into a customer level.

For housing loan, LGD historical constant model was implemented as Champion model. While, multinomial LGD model was selected as Challenger model.

## Data Management

### Overview of Input Data Set

The input for actual LGD calculation includes outstanding balance at default, state of default, final flag and actual cumulative recovery amount. The sum of the present value (PV) of these cash flows amounted to the accumulated recovery amount which is then used to calculate the value of LGD. The state of default indicates a posterior state of action after the customer has defaulted. The final flag indicates whether there is an on-going collection attempt and thus whether the state indicated are final. In summary, the data requirement for the calculation of actual LGD is shown as follows:

Figure 14, Data required for actual LGD calculation

### Sample Design

All data of home loan customer in RI, write-off and collateral sale were for the period between 2002 and 2018. For LGD model development, this model will focus only on NPL customers without any identified data issue which can be classified into 2 groups:

Finished cash flow process customers which recovery can be calculated by sum of discount cash flows divided by outstanding at NPL date.

Unfinished cash flow process customers which recovery can be calculated by sum of discount cash flows plus expected recovery divided by outstanding at NPL date.

### Data Cleansing and Exception Handling

When the modeler investigates recovery cash flow of each customer, data issues are found and data cleansing is required. After discussing with related parties, the modeler found that net recovery cash flow is unreliable.

Multiple same amount of net recovery cash flow: The multiple amounts are the exact amount of sale price or appraisal value of collateral. This leads to overestimated recovery and LGD becomes smaller than it should be.

Net recovery cash flow in wrong period: Net recovery cash flow is recorded before or after the event. For example, collateral is sold at May 2016 but recovery cash flow is recorded at Jan 2016. The difference in time period would impact the concept of time value of money.

Missing cash flow: although collateral is sold, no recovery is recorded.

Disharmony in recovery cash among three data sources: For example, outstanding at default is 100. Net recovery cash from RI is 40. Write-off amount from write-off is 30, meaning that customers pay back 70. However, recovery cash from collateral sale is 80. Based on the same customer, these three sources show different recoveries.

To deal with problem 1 – 3, the modeler needs to clean all wrong net recovery cash flow and place the right one in the right period. For problem 4, after discussing with AQ department, the rank of priority is collateral sale > write-off > RI. It means that if recovery cash flow from three sources is conflict, the modeler believe in collateral sale, then in write-off and finally in RI.

### Final Development Sample

The data was filtered in the process of data cleansing to exclude data issues stated in 6.4.3. As a result, the total number of 14,146 NPL customers will therefore be used in LGD calculation.

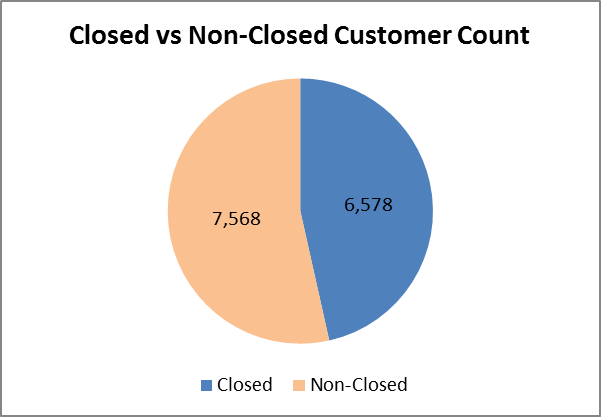


Figure 15 Number of customer with closed files and non-closed files

From above Figure 15, the final data set used in this LGD model contain 6,578 closed files customers which is 53.3% of total and 7,568 non-closed files which is 46.5% of total.

## Model Development

### TTC Model

The LGD historical constant methodology is based on the discounted of actual cash flows that can be recovered by the collection process from the date of default to the end of the recovery process. In addition, BIS (2004) states that banks who choose to calculate realized LGD using the workout method must include the direct and indirect costs associated with the collection of the exposure. Thus, we consider both costs in our LGD calculation. Direct costs are those associated with a particular asset, including fees for an appraisal of collateral, costs of selling assets, costs of running a business, and other professional fees. Indirect costs are necessary to carry out the recovery process but are not associated with individual facilities. For KBank portfolio, collateral is generally held against a customer, not against an individual loan. Thus, for the calculation of historical LGD, it makes sense to view all of a given customer’s facilities as a single exposure, and calculate the severity based on this.

**Non-Closed Files Customer Treatment**

When using the workout method, the problem arises of how to deal with partial recovery profiles of non-closed files. The non-closed files refer to defaulted customers in the database who do not yet complete the loan recovery process (i.e. many of the loans are still in the process of debt collection).

The simplest approach is to exclude these non-closed files from the LGD estimation process, with LGD based on closed files only. Whilst simple, results based on this approach may be affected by data selection bias if the non-closed files contain information relevant to LGD which is not captured by the recovery profiles of the closed files. Moreover, inclusion of the non-closed files may still be relevant if they contribute to reduce the error around the estimates (Rapisarda and Echeverry, 2010).

As a result, we include the non-closed files in the LGD estimation process and need to make an assumption on the recovery rate of the non-closed files. In doing so, we first divide the closed files into four main default pathways, as shown in

Table 38, and estimate each group recovery rates:

|  |  |
| --- | --- |
| Path | Explanation |
| Self Cured | The defaulted customer takes a certain amount of time to recover. However, No significant loss and no change in the structure or conditions of the facilities. |
| Early Cured | The defaulted customer takes a certain amount of time to recover and might not be able to fulfill his/her contractual obligations from time to time. However, No significant loss and no change in the structure or conditions of the facilities. |
| Restructuring | The defaulted customer recovers after a restructuring of his/her facilities. Usage of collateral may sometimes be part of the restructuring. Loss amount varies whereas customer relationship maintained |
| Liquidation | All facilities of the defaulted customers are liquidated, i.e. sales of loans, usage of collateral, etc. Loss amount is generally higher than that observed from restructuring. End of customer relationship. |

Table 38, all possible paths and their definition

However, Analysis of characteristics and historical LGD between self-cured and early cured is very similar. Figure 16 shows that the recovery curve of closed accounts in path of self-cured and early cured is slightly different in early period but is very close in later period.

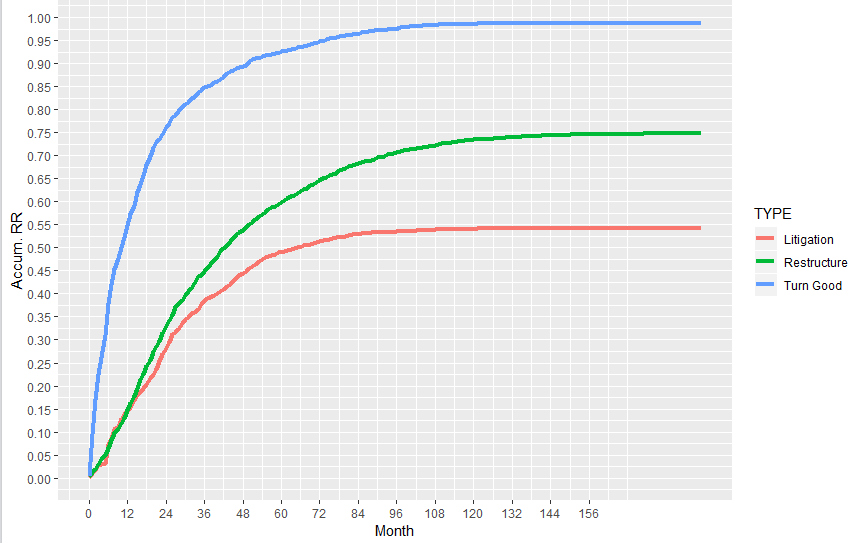


Figure 16 Recovery rate of closed accounts by path

The evidence supports that these two groups, self-cured and early cured, is very similar. The modeler decides to group them together as shown in table below.

|  |  |
| --- | --- |
| Path | Explanation |
| Early Cured | The defaulted customer takes a certain amount of time to recover and might not be able to fulfill his/her contractual obligations from time to time. However, No significant loss and no change in the structure or conditions of the facilities. |
| Restructuring | The defaulted customer recovers after a restructuring of his/her facilities. Usage of collateral may sometimes be part of the restructuring. Loss amount varies whereas customer relationship maintained |
| Liquidation | All facilities of the defaulted customers are liquidated, i.e. sales of loans, usage of collateral, etc. Loss amount is generally higher than that observed from restructuring. End of customer relationship. |

We then classify the non-closed customers into each group of state and project each group’s recovery rates following those of the closed files. We project the remaining cash flows of non-closed customers based on the average recovery rate of closed group of customers by month. Here is the equation for this approach:

Expected Recovery (%) is derived from projected recovery rate curve of each state and segment group. We can generate this curve by dividing all cash flow by outstanding of all closed-customers in that particular group and repeat this step for the next month till the end.

For example, if the customer has % Actual Recovery equal to 20% at month 36 after NPL date. Then we get an expected recovery from the curve which appears to be 30% from month 36 to the end of recovery curve. As a result, % Recovery of this customer must be 50% (20%+30%). For LGD, the following formula will be used:

**Final Model**

Actual recovery rate of closed files customers are shown below.

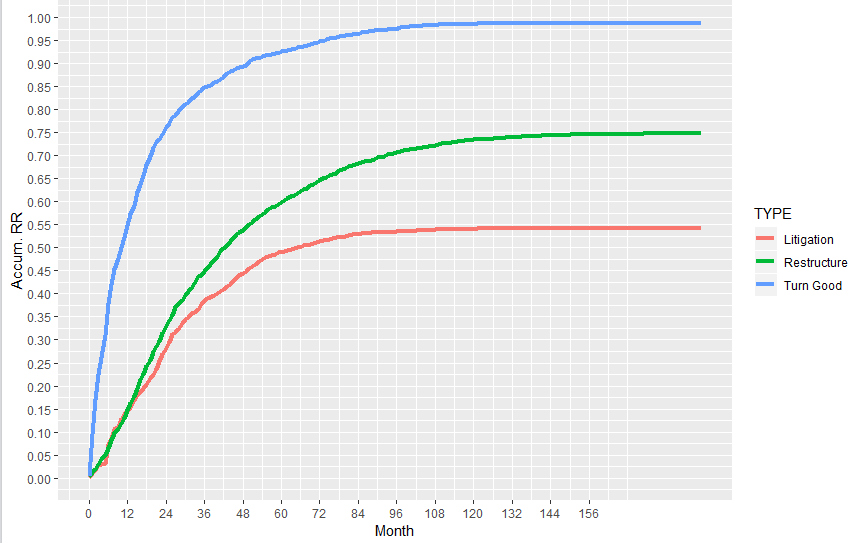


Figure 17 Recovery Rate of closed files customers

From the figure 15, the Accumulate Recovery Rate of Turn Good, Restructure and Liquidation Customers are 96.53%, 73.76% and 47.33% respectively. In term of flat period, turn-good customer is in 158th month on book, Litigation is in 171st month on book, and Restructure is in 190th month on book.

After combining the projected recovery rate of non-closed files customer (refer to the methodology mentioned above) and closed files customer, the results of recovery rate of each path are as following:

Figure 18 Recovery Rate of turn-good customers

Figure 19 Recovery Rate of restructure customers

Figure 20 Recovery Rate of litigation customers

To summarize, the weighted average LGD by outstanding of a whole HL portfolio is 33.70%. LGD for Turn-good, Restructure and Litigation customers are 14.18%, 31.11% and 47.69% respectively.

However, due to KBank propose to use a constant at LGD at 34.42% for TFRS9 which is greater than current actual observed weighted average LGD. So KBank decided to use a constant at LGD at 34.42% for home loan customers in TFRS9.

### PIT Model

Currently, there is no model for point-in-time LGD.

## Pre-Validation

### TTC Model

**Validation Methodology**

R-squared (R2) is a statistical measure of how close the data fitted regression line. The larger R2, the better model it is. Kbank compares R-squared (R2), Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) of training sample and testing sample to ensure performance and robustness of the model.

Mean squared error (MSE) will be used as the formula shown below. The smallest MSE value, the better model it is. The validation result is shown in table 6.

For the LGD back testing, key Assumption for LGD model is past experience can be used to predict the future (i.e. the Bank’s loss in the future would be similar to ones of the past given customers that have similar characteristic). Then, a one-side binomial test at 95% confidence interval will be test. The binomial test is done by comparing the predicted LGD values of the portfolio with the realized (actual) loss from the database. Since not all of bad customers in database finished the loan recovery process (many of the loans are still in the process of debt collection), we will need to make assumption on the recovery rate of those unfinished as well.

**Validation Result**

The weighted average LGD by outstanding of a whole HL portfolio is 33.70%. LGD for Turn-good, Restructure and Litigation customers are 14.18%, 31.11% and 47.69% respectively.

To sum up, the average actual LGD of the validation sample is 33.70% which is less than 34.42% so we can conclude that our constant LGD can cover the average actual LGD for housing loan portfolio as shows table below;

|  |  |  |
| --- | --- | --- |
| Product | Observed  LGD | LGD |
| Housing Loan | 33.70% | 34.40% |

# Criteria for a Significant Increase in Credit Risk (SICR)

With the new IFRS9 loan loss provision, loans are classified in three stages: stage 1 – initial recognition (yet to be impaired), stage 2 – significant increase in credit risk, and stage 3 – objective indicators of impairment. For loans in stage 1, banks need to estimate 1-year expected credit losses. On the other hand for loans in stage 2 and 3, banks need to provide provision and thus estimate expected credit losses for the whole lifetime of the loans. Both the 1-year and lifetime expected credit losses estimation shall reflect the banks’ forward looking macro-economic view.

The criteria for significant increase in credit risk (stage transfer criteria) will be used for the classification of loans between stage 1 and stage 2. KBank employed both the qualitative and the quantitative criteria for stage classification. An example of qualitative stage transfer criteria stage is the 30+ days past due (DPD), fraud, black list from the revenue department and so on. Thus, all instruments with DPD>30 days are automatically classified as stage 2. This model development document solely focus on the formulation of the quantitative stage transfer criteria.

This section of the document outlines the methodology review and the development including any expert opinions and judgements of our champion significant increase in credit risk criteria. KBank explore four methodologies to develop the quantitative criteria the significant increase in credit risk namely the rating downgrade criteria, the remaining lifetime PD criteria, the forward run test and the high credit risk region. In each of the methodology, KBank leverages existing behavioral scorecard, supermaster rating transition matrix and PD term structure to induce statistical inference and the formulation of the criteria itself.

## Scope

Both the qualitative and quantitative stage transfer criteria are applied to all non-NPL instruments. All non-performing loan instruments are automatically assigned to stage 3.

## Modelling Methodology Review

### Rating Downgrade

The rating downgrade is a quantitative stage transfer criteria that contemplate on the rating downgrade compared to the rating assigned at the loan origination. If the number of rating downgraded of an instrument is greater than a given threshold then the instrument is deemed to be a significant increase in credit risk and assigned to stage 2. The number of rating downgrade can be identified illustratively in the figure below.

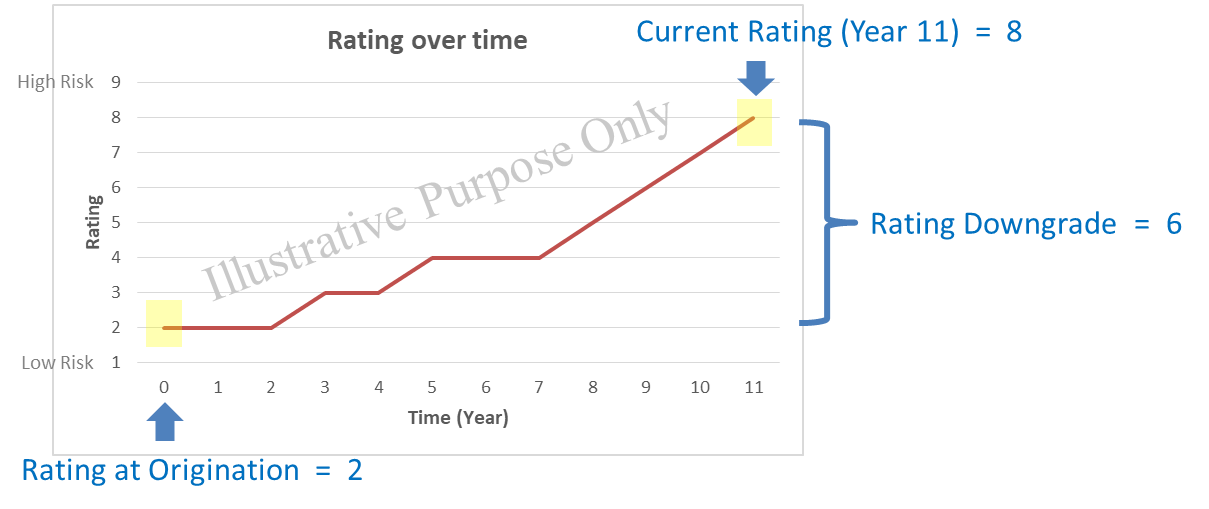


Figure 21, Rating observation: origination vs. reporting

For retail portfolio, we use an empirical rating migration from bootstrapped samples to construct the threshold for notch change (between 75-85 percentiles depending on the portfolio).

### Remaining Lifetime Probability of Default

The remaining Lifetime PD stage transfer criteria recognize if the remaining lifetime probability of default at the reporting date has deviated statistically from the lifetime probability of default estimated at origination. This measure is highly sensitive to our choice of PD term structure model. By default this compares the lifetime PD generated from the champion model of each portfolio.

In short, the significant increase in credit risk under this criteria is SICR is reflected by statistical deviation in [1] the remaining Lifetime PD estimated at origination compared to [2] Lifetime PD at reporting date. The statistical deviation in this criteria is founded on a confidence band based on the average deviations in the marginal probabilities over time (for each supermaster rating). The standard deviation for our confidence band is illustratively measured as per the following figure

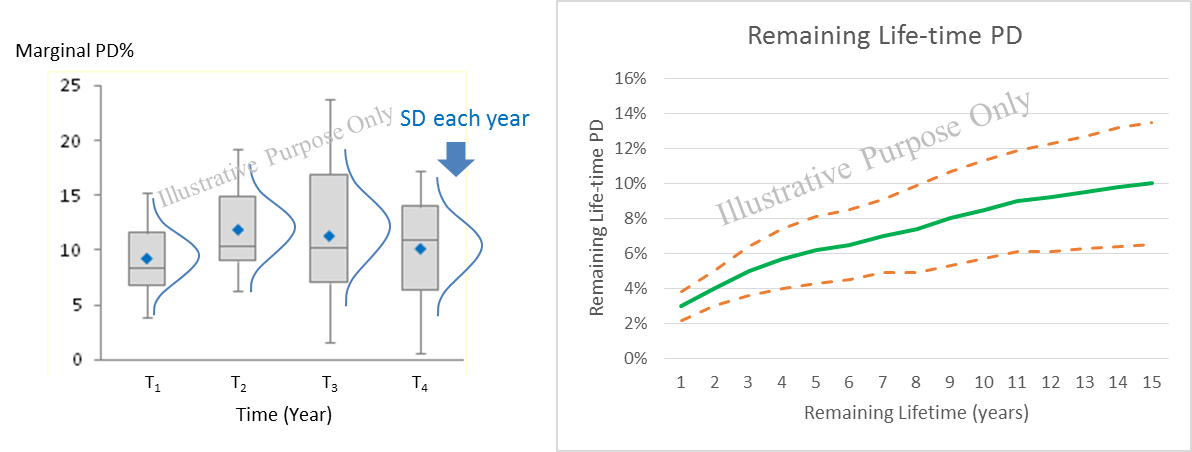


Figure 22, Standard deviation and confidence band for Lifetime PD

After the average standard deviation of marginal probabilities has been defined . We generate the confidence band for each of the remaining lifetime as follow

Where is the average standard deviation of the mariginal PD for rating

is the cumulative PD at estimated origination (origination rating ) upto maturity date

is z-value of a confindence level (by default 80%, two-tailed)

### Forward Run Test

The forward run test criteria check if the Forward PD at reporting date has significantly increased from the Forward PD at the similar point of time estimated at origination. If we were to plot forward PDs of two time series [1] Forward PD from origination and [2] Forward PD at reporting date, we can compare the two time series using run test to see if they are statistically different as illustrated in the figure below.

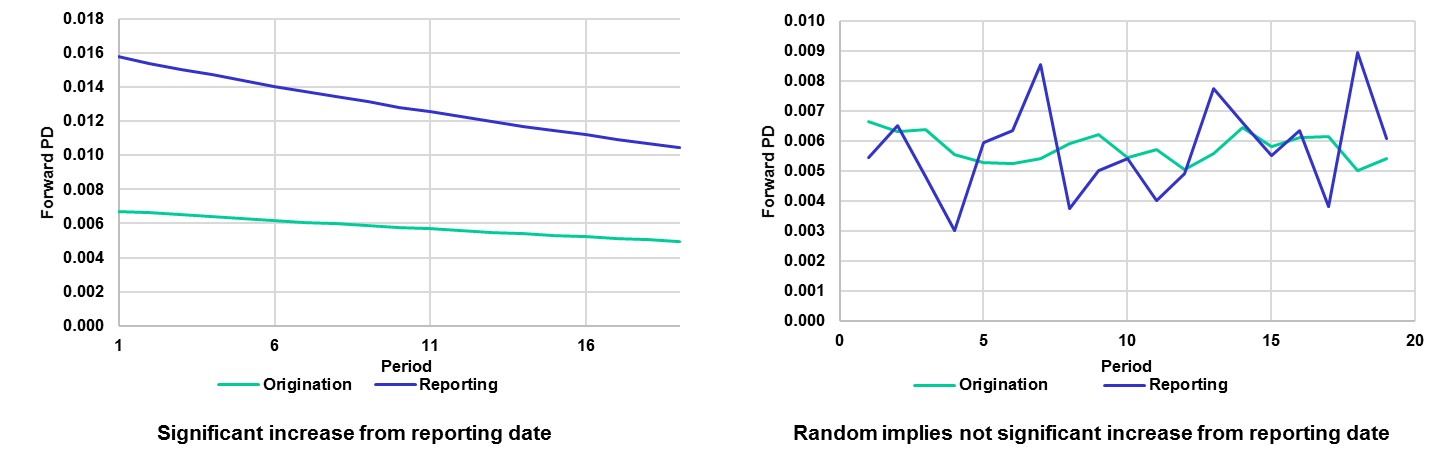


Figure 23, forward PD run test

The hypothesis of the forward PD run test is defined as follow

*H0: Forward PDs at origination date and at reporting date are not different*

Test statistic

Where is the observed number of runs

are the number of origination and reporting date values

### High Credit Risk Region

The high credit risk region is our final significant increase in credit risk criteria which may be use in conjunction with other criteria. The high credit risk region is defined as a supermaster rating threshold. If a credit rating of an instrument are greater than the prespecified value then it is automatically assigned to stage-2.

The conception of the high credit risk region would start with an observation of the average rating of a portfolio. We’ll then use an associated rating downgrade threshold as the high credit risk region for the portfolio. The setting is further validated by testing the equivalent of survival curves i.e. between the low credit risk and high credit risk region.

The hypothesis of the equivalency of survival curve is defined as follow

*H0 : The customers with credit rating higher than the HCRR threshold have equivalent survival curve to those customers with ratings better than the threshold.*

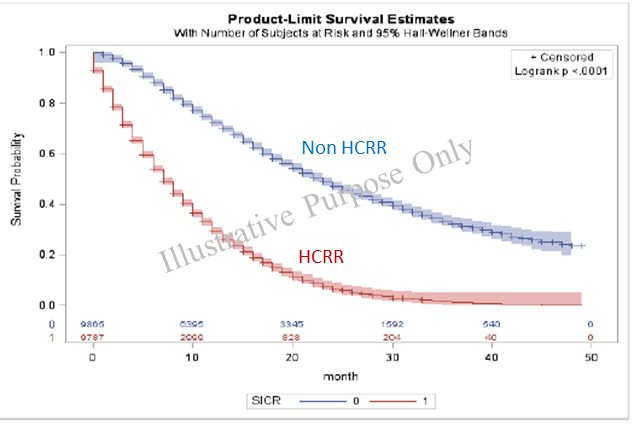


Figure 24, Equivalency of survival curves between HCRR and non-HCRR

## Model Development Approach

For significant increase in credit risk criteria, KBank believe that the rating downgrade and the high credit risk region are the appropriate methodologies to be used as a chamption criteria. The main reason for this is the fact that the methodology is lower in complexity in terms of both the data requirement and system implementation. The inputs of the model and the development exercise aligns with the bank’s current data structure. The output of model also aligns with current business model and TFRS9 risk requirements in term of classification of instruments whose credit risk have increase significantly. Other methodologies are explored and implemented as challenger SICR criteria.

This section briefly outline the steps that have been taken in the development of the champion SICR criteria i.e. rating downgrade and high credit risk region. The development SICR model is done successively to the development of PD model and will be using an identical sampling data. Figure below outline the steps taken in the development of rating downgrade and high credit risk region.

Figure 25, Steps of development: stage transfer criteria

.

The development process starts with consolidation of portfolio data across time into a single data set (these are the same data set which are used in the PD model development). Then development samples are created by bootstrap sampling from the large data set. We then observe the 75 to 85 percentile of migration for each of the initial rating in each of the sample. The observed mode of the 75-85 migration percentile (based on expert opinion) across multiple bootstrapped samples are then selected as a downgrade criteria.

For the high credit risk region, we observe the average supermaster rating of an entire portfolio in the data set then we use the downgrade criteria for that rating as a high credit risk region.

## Data Management

For data management, please refer to the Data Management section for the modelling of the probability of default term structure.

## Model Development

### Rating Downgrade

The development of rating downgrade criteria starts with the bootstrapped sampling to observe the rating (supermaster) migrations. For each initial rating, we then observe the distribution of the 12-month rating change as per the following figure.

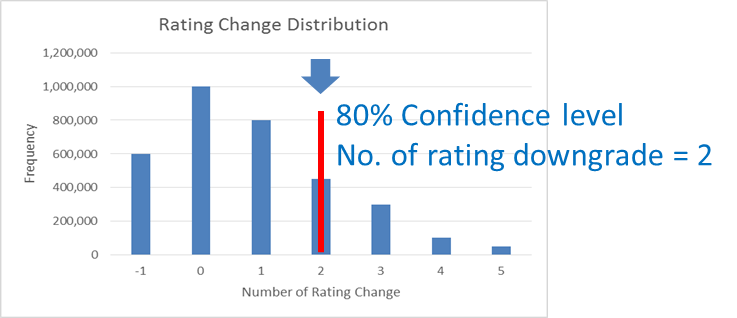


Figure 26

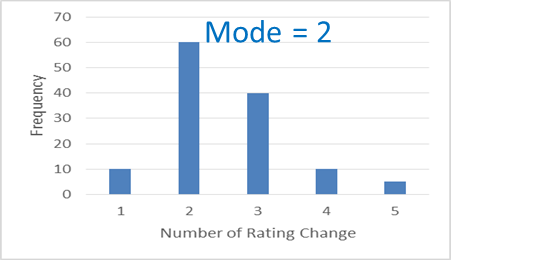


Figure 27, Rating change distribution and mode of bootstrapped sample

The steps are repeated across multiple (100) bootstrapped samples to observe (100) of 80th percentile of the 12-month rating change. The mode of these 80th percentile numbers is then chosen as a downgrade criteria. The result of this exercise are the number of downgrade rating change for each of the initial (origination) rating. The result is then overridden with few expert judgement to maintian the monotonicity of the downgrade for each of the initial rating. The results are as follow

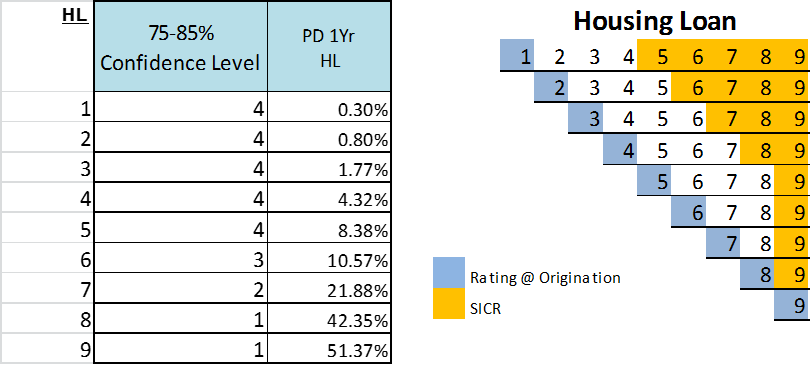


Figure 28, Final rating downgrade criteria

### High Credit Risk Region

As mentioned previously, the high credit risk region is essentially the rating downgrade (change) of the average portfolio rating which in this case is rating 4. The high credit risk region is thus rating 8 or greater. The distribution of customer that falls under and in the high credit risk region are shown in the table below.

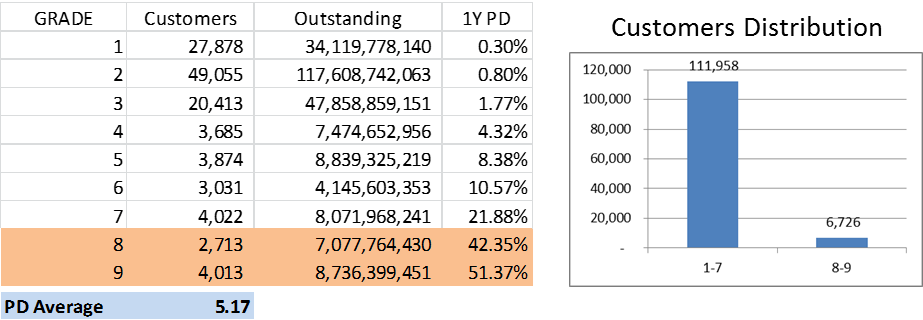


Figure 29, Final high credit risk region criteria

We validate this setting by considering the survival curves of each segment i.e. low credit risk region vs. high credit risk region. The validation results show that the survival curves are vastly different and the high credit risk region setting is appropriate. Results of validation are shown below.



Figure 30, Validation of HCRR, equivalency of survival curves

1. http://www.IFRS.org/current-projects/iasb-projects/financial-instruments-a-replacement-of-ias-39-financial-instruments-recognitio/documents/IFRS-9-project-summary-july-2014.pdf [↑](#footnote-ref-1)