Model Development Documentation

Applicable to: Retail Loss Forecasting & PPNR Models

Note on Model Risk Management – Document Automation

With the implementation of "Model Risk Management – Document Automation", on the user's instruction, the system will generate automatically the documents used in Model Risk Management (MRM) for Validation, Annual Model Review and On-going performance monitoring, before the appropriate workflow is submitted.

The automation process will merge information from the BankRisk Model Risk Management System (BankRisk MRMS) and the content of existing files (the uploaded template such as MDDT, MVRT, AMR Sponsor, OPA Sponsor etc.), to create a pdf file (Automated Document). This is to ensure data consistency between inventory in the BankRisk Model Risk Management System, and MRM documents (Development, Validation, Approval, and On-going performance monitoring).

All Automated Documents (Sponsor) will have the same general structure with the following sections.

- *Cover Page*
- *Table of Contents*
- *BankRisk MRMS Sections:*
  - *A.1. Model Profile*
  - *A.2. Model Risk Rating*
  - *A.3. Inputs, Outputs & Assumptions*
  - *A.4. Limitations*
  - *A.5. Exceptions*
  - *A.6. Document body*

Model Info Sheet

iMRMS Model ID(s)	167125
Model Name	NA Mortgage Method A - Residential Mortgage Model
PMACS ID & Description (for Risk Models Only)	13587774, 13587771, 1611468
Region	North America
Country	USA
Product(s)	Non-Modified Residential Mortgage
Model Usage	See Model Usage Grid in Section 1.1
Model Materiality Tier (MMT)	In iMRMS, Pre-MMT is High
Custom/Vendor	Custom
Consumer/Commercial/PPNR	Consumer
Estimation Technique	Competing Risk Transition
Model Output	See Model Usage Grid in Section 1.1
Number of Segments	PD Model – 28 Modeled & 28 Non-Modeled Segments LGD Model – 6 equations
Delinquency Bucket	All Delinquency Buckets
Model Implementation System	Consumer Analytics Grid (SAS Grid)
Planned Implementation Date	January 2019
Target First Use Date	March 2019
Model Sponsor	Fred Bader– Executive Model Sponsor NA Mortgage Chief Risk Officer
Model Sponsor (delegate)	Subodh Rai - Consumer Analytics Modeling Unit (CAMU) Managing Director
Model Developer	Bin Sun, CAMU Director
MDD Submission Date	9/30/2018; 11/21/18; 12/21/18
Person(s) Completing MDD	Indranee Saha / Ting Sun / Boying Liu / Jie Chen / Ying Li / Su Ge/ Choi Lin Lai / Jonathan Franco / Jixiong Yu / Jie Tang / Shiping Wang / Shuangxi Xie / Sriram Popuri / Juan Zhao / Jennifer Conner / Bernadette Lato / Sri Manjunatha / Frank Oglesby

Document Version Control

Version	Date	Author	Description of Change
1	09/30/2018	CAMU	Initial MDD submitted for MEA
2	11/21/2018	CAMU	First resubmission to address MEA feedback
3	12/21/2018	CAMU	Second resubmission to address MEA feedback

Portfolio Info Sheet

	Portfolio Description	Product	Portfolio ID[1]	Open Accounts	ENR (US $M)	Annual NCL (US $M)	NCL %	Date (as of)
1	Bank Mortgage - Residential Corp	Mortgage	13587774, 13587771	95,051	$40,334.64	$6.31	0.02%	6/30/2018
2	Bank Mortgage - Residential Holdings	Mortgage	1611468	68,164	$7,570.05	$5.60	0.07%	6/30/2018
3	Bank Private Bank Residential Mortgage	Mortgage	N/A	22,411	$26,516	$2.29	0.01%	6/30/2018
4	Serviced For Others – Corp	Mortgage	N/A	240,818	$45,681.69	NM[2]	NM[2]	6/30/2018
5	Serviced For Others – Holdings	Mortgage	N/A	94,010	$11,563.29	NM[2]	NM[2]	6/30/2018

[1] *Portfolio information here comprehensive and not specific to only modified loans as these are PMACS level details*

[2] *The Ending Net Receivable (ENR) of the Serviced For Others portfolio is not receivable by Bank. These loans have been sold to investors, and only the servicing rights have been retained. There is no NCL as MSR exposures are off balance sheet.*

Model Development Documentation Template (MDDT) Guiding Principles

The Model Development Documentation Template (MDDT) ensures consistency and replicability in the model development and documentation process by providing clear guidance on what is needed to assess model eligibility and perform model validation.

The Model Sponsor and Model Developers are responsible for substantiating, documenting and justifying all decisions and choices made during the modeling process, including performing testing required for substantiation.

Model Validators will evaluate the reasonableness of the decisions and choices made based on the documentation provided by the Model Sponsor (and Model Developers).

Because the same MDD template is used for all types of Modeling Approaches (e.g., statistical, judgmental, multi-step), not all questions will apply to all approaches (e.g., questions on statistical tests on historical data do not apply to Modeling Approaches that lack historical data.)

Executive Summary

[Please provide a high-level summary of the findings and contents of the document, including

- Background of the model and history,
- Usage of the model (for e.g. CCAR, CECL, ICAAP, IFRS9, Business Planning etc.),
- Regulatory requirements,
- Bank's/Sponsors approach to meet the regulatory requirement,
- Discussion on why the same model can be applied to multiple usages and
- Changes since previous validation.

Provide narrative on the model chronology / evolution that covers a high level summary of model updates, changes, and enhancements etc. prior to this model arrival.

Include a non-technical description of the model. Describe the model framework in easily understood terms along with a model process flow diagram, from input to output. Provide an assessment of Model Performance. Provide model output from the final model specification, along with rationale from model developers for reasonableness of results. Address results of testing and the overall soundness of the methodology. Describe the regulatory feedback regarding the model, its previous versions, related models and portfolio. Explain how the feedback was addressed in the current model].

This document should comply with the expectations of Bank's internal Model Risk Management policies, policy supplements, procedures, standards and guidance, as well as meet expectations from relevant regulatory guidance and related documents, including:

- **FRB SR 11-7**: "Supervisory Guidance on Model Risk Management", April 2011/ OCC Bulletin 2011-12: "Supervisory Guidance on Model Risk Management", April 2011;
- **FRB SR 15-18**: "Federal Reserve Supervisory Assessment of Capital Planning and Positions for LISCC Firms and Large and Complex Firms", December 2015;
- **CCAR 2015 Summary Instructions**: "Comprehensive Capital Analysis and Review 2015 Summary Instructions and Guidance", October 2014;
- **FR Y14-A**: Appendix A: Supporting Documentation in "Instructions for the Capital Assessments and Stress Testing Information Collection (Reporting Form FR Y-14A)", March 2015.
- **CECL, IFRS9 regulatory standards/guidelines.**
- **MODEL RISK MANAGEMENT, Policy & other Supplements -** Model Testing Guidance, Gating Principles, Code & Data Guidance]

☐Submission includes a model change (Please mark if applicable):

Due to the extensive framework redesign, CAMU deems that it will not be sufficient to consider Method A redevelopment efforts as a model change. Instead, CCAR 2019 Method A suite is considered an overhaul of the CCAR 2018 version, with newly created model ID.

☐Changes in the model specification (e.g., inputs, processing components and techniques, outputs)

☐Changes to model use (e.g., using an existing Model for a new portfolio, geography or segment)

[If submission includes a model change, provide a detailed description of the proposed change in Section 9.4 Model Changes]

Abstract

The Federal Reserve's annual CCAR is an annual exercise that assesses the capital adequacy of large, complex U.S. bank holding companies (BHCs), given their unique idiosyncratic risks. The Dodd-Frank Act Stress Test (DFAST) submission is a semi-annual exercise that caters to the same core purposes as the CCAR, mostly serving as critical means for supervisory oversight and enforcing tangible actions. As regulatory stress testing techniques have already been vetted as effective tools to assess the robustness of the banking structure, the meticulous development of a credit risk model is of paramount importance for all risk management activities.

This Model Development Template (MDD) describes the re-development and testing of the NA Mortgage Method A Residential Mortgage (henceforth Method A RM) model for use in the annual CCAR/DFAST cycle for all FRB and BHC econometric scenarios, mid-year CCAR cycle BHC scenarios. The entire model suite is composed of a series of loan-level econometric models that are related through common dependence on macroeconomic as well as loan and borrower specific factors. Using these loan-level models, the model can forecast the delinquency, default, prepayment and loss severities for the North America real estate portfolios across various economic factors and scenarios.

The 2019 Method A RM model results demonstrate a robust framework capable of predicting both the stress and recent period losses with increased predictive accuracy across model backtesting and sensitivity analyses. The primary factors behind the model's improved performance are due to the inclusion of more recent period data that captured the current trends of portfolio composition mix reflective of stronger credit profile in recent times, quantitative evaluation of the model's segmentation scheme and carefully selected variable transformations leading to better cohort level performance, systematic segregation of the number of modeled vs non-modeled equations backed by statistical analyses and improved valuation of distressed properties and severity model.

Model Usages

As mentioned earlier, this model can be used to forecast losses for all five CCAR macro-economic scenarios which include the BHC baseline, BHC stress, FRB baseline, FRB adverse and the FRB severely adverse.

Apart from meeting the CCAR specific mandates, the model output can also be used for Non-CCAR initiatives, which include the Allowances for Loan for Lease Losses, Loan loss Reserve and Troubled Debt Restructuring processes (ALLL, LLR and TDR). The purpose of the ALLL/LLR/TDR is to reflect estimated credit losses/provision within a bank's portfolio of all its loans and leases. This provision is used to cover a number of factors associated with potential loan losses including bad loans, customer defaults and renegotiated terms of a loan that incur lower than previously estimated payments. By setting aside loan loss reserves and constantly updating estimates through loan loss provisions, banks can ensure they are presenting an accurate assessment of their overall financial position. This financial position is often released publicly through the bank's quarterly financial statements and is considered a critical element of regulatory assessment of bank's overall liquidity and health.

The model is also used for Business Planning purposes to develop a business direction (sales, fulfillment, servicing) to achieve the desired level of earnings or return target, and as such a lot of the business decisions are partially based on the NCL/Delq/LLR forecasts that are generated by the model suite. Further, the outputs of this model also serve as key inputs for managing Bank's internal Risk Appetite Framework (RAF) strategy. Risk appetite is considerably more than a sophisticated key performance indicator (KPI) system for risk management. It is considered one of the core instruments used for better aligning overall corporate strategy, capital allocation, and associated riskiness. For additional details on the RAF, please refer to the Appendix section.

In addition, the model is also proposed to be used for CCAR and non CCAR evaluation on the 'Serviced for Others' (SFO) loan portfolio. Model performance testing results for the SFO loan portfolio is presented in Chapter 6.

The model is capable of producing month-over month losses and delinquency projections (units and balances) across time to support regulatory requirements as well as in-house business planning initiatives such as allowance for loan, loss and leases (ALLL), troubled debt refinancing (TDR) and portfolio risk management (risk appetite).

Background

The Method A RM model is a suite of models; the probability of default (PD) model which is itself a collection of models yielding a monthly transition matrix, EAD which is based on loan amortization schedule, and the loss given default (LGD) model which models the loss severity rate.

The PD model uses a competing risk transition matrix approach to calculate loan level monthly transition probabilities for various intermediate delinquency states, involuntary payoff (IVP) as in loan foreclosure or bankruptcy and voluntary payoff (VP) such as prepay/refinance, as function of loan-specific and macroeconomic factors. The transition matrix model is applied over the forecast horizon at each month to compute the probability of transition to the next state (terminal – VP/IVP and non-terminal-transitions to worse, cure or partial cure). The PD model uses a probabilistic approach to arrive at the final probabilities. While the PD modeling framework has remained the similar from prior year's version, the LGD modeling approach was redeveloped for the first lien loans. While the prior year's LGD model modeled losses by their disposition types, the 2019 LGD framework models losses by their loss type (Full vs Partial vs Zero losses). This not only allowed the alignment of the loss methodology across liens but helped remediate the prior model's shortcomings around exclusion of zero losses and the non-monotonic relationship to marked-to-market LTV which is considered a critical input of the Severity model. Both the PD and the redeveloped LGD models, when combined with business logic/algorithms, work together in a system to yield economically important metrics, resulting in forecasted default and prepay balances and units along with the losses (gross and net). The loan amortization module was also enhanced this year as part of ongoing production code improvement efforts.

Model Scope

The Method A RM model is built to cater to the US real estate portfolio that includes the two specific lines of business - BankMortgage, Inc. (CMI), Bank Private Bank (CPB). The residential mortgage loans in scope include never modified mortgage loans (across all liens) held in portfolio as well as those Serviced for Others (SFO) loans. Please note while the model is developed solely on the held in portfolio loans (for additional rationale and justification for not having a separate SFO model, please refer to Section 1.1 of the MDD), it can be borrowed for modeling the SFO loans too. The model's performance for the SFO portfolio has been illustrated in Chapter 6 of the MDD.

US Economy at a Glance

Real estate represents a significant portion of most people's wealth, and this is especially true for many homeowners residing in the United States. According to the US Census Bureau, the number of households in the U.S. has been growing steadily over the past decades, as has the population. The total number of households has doubled from about 63 million in 1970 to more than 126 million in 2017. The real estate market is typically influenced by the state of the economy, generally measured by economic indicators such as gross domestic product (GDP), employment data, income, interest rates, stock market movements measured by S&P500 Index and the housing price index (HPI). All of these factors affect the real estate performance either directly or indirectly through changes in borrower/investor behavior and hence are considered critical when evaluating the credit risks of the US real estate portfolios. Presented below is a quick peek at the state of US Economy as of 2018Q1.

- Housing Price Index (HPI) - The HPI indicator provides the trend in housing prices over time that heavily influences the changes in the rates of mortgage defaults, prepayments and housing affordability. The HPI witnessed an increase of 9% from quarter four of 2017 to quarter one of 2018.

- Interest Rate - The mortgage rates are typically tied to factors like the returns on 10-year Treasury note. The 10 yr. Treasury yield correspondingly rose to 2.75% in quarter one of 2018 from 2.5% end of Q4 2017.
- Unemployment - Unemployment rate remained steady at 4.1% for the first quarter of 2018, well above the Fed's long-term equilibrium estimate, from the quarter before.
- GDP – The GDP measure the overall value of economic activity. GDP grew by 2.0%, the first quarter of 2018 from Q42017
- Income – State level aggregated household income is the summation earned by all factors of production in the form of rent, wages, interest and profit.
- S&P500 - The S&P 500 is widely regarded as the single best gauge for evaluating the overall health of the US equity market. The S&P500 index has stayed stable from quarter four of 2017 (2673.61 - Close price) to quarter one of 2018 (2,648.05 – Close Price) reflecting the ongoing strong state of the economy.

Portfolio Description

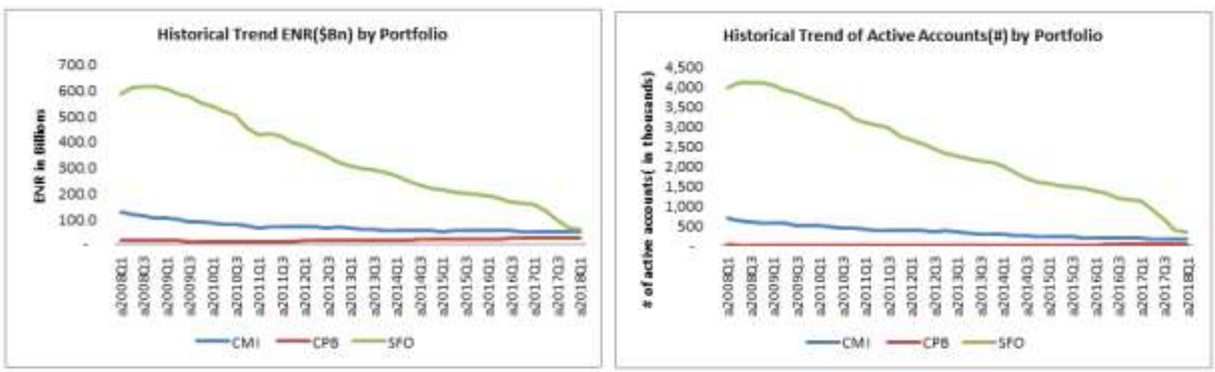
The RM portfolio includes loans from the BankMortgage (CMI) and Bank Private Bank (CPB) businesses.

The CMI entity is currently the largest carrier of residential mortgages for Bank's North America Mortgage business with the total size for the portfolio approximated at $47.1Bn in ending net receivables (ENR), as of the first quarter of 2018. It is important to that that this portfolio has shrunk considerably over time from $125Bn (2008Q1) in managed assets at the onset of the crisis period to almost one-third of its former volume in current period. The portfolio delinquencies (90+DPD) and the credit losses (NCL) correspondingly have also gone down considerably from $3.9Bn and $134MM from the time of the crisis [2008Q1] to $0.67Bn and $6.5MM dollars respectively as of 2018'Q1. The notable drop in portfolio size and corresponding delinquencies & losses was strongly driven by the management's desire to improve portfolio mix post the crisis experience with an onward shift towards more retail originations compared to correspondent/broker based originations, with a focus on an incentivized relationship pricing model, stringent underwriting criteria and enhanced digital capabilities.

The CPB business on the other hand has grown considerably from $17Bn in 2008Q1 to $25Bn in 2018Q1 and continues to grow in the near future, given the strategic approach to cater to private bankers (Law Firm Group, Ultra Net Worth, High Net Worth) with home financing needs for primary residences, vacation homes and investment properties, through a wide range of loans products, structure and unique ownership entities. The portfolio has witnessed sparse but volatile net losses (NCL) over time that is reflective of affluent customer bases that are usually not sensitive to business cycle movements.

An important point worth mentioning here is the applicability of this model suite towards estimating losses for BankMortgage's Serviced for Other (SFO) portfolio. The SFO portfolio reflects the sold and serviced volume that Bank holds on to its balance sheet. As of Q2-2017, BankMortgage made the decision to exit the Servicing business in a gradual phased out manner. The decision was based on a forward-looking view to create a leaner organization with a focus towards integrating retail-banking operations with mortgage originations that would better position the business for the future. Given this strategic business decision, the size of the SFO portfolio has rapidly shrunk over the last fifteen months. This diminishing portfolio trend has been evidenced in the chart below. Currently the SFO portfolio has approximately $58.3Bn in unpaid principal balances as of quarter one of 2018.

Portfolio Trends (2008Q1 – 2018Q1)



Model Performance Assessment

As part of the development and subsequent testing process, the model performance was assessed through back testing across various time period horizons as well as sensitivity analyses. All model results have been evaluated based on Model Risk Management (MRM)'s Guiding Principles and Loss Forecasting Model Testing Guidance, with the following key highlights. Please refer to Chapter 6 for all Model Performance related details.

> The 2019 Method A RM model has demonstrated better predictive accuracy in both recent and stress periods on all backtesting horizons compared to the existing model. The model performs satisfactorily over the twenty-seven month testing horizon for the CMI portfolio as illustrated in Table 1 below. CPB portfolio(Table 2) continues to produce conservative estimates due to overall lack of CPB delinquency/loss data over the entire model development period and the erratic (volume and volatility) trend of losses during the historical stress period. Please note that

for CPB portfolio, the actual net credit losses don't meet Model Risk Management's (MRM) Materiality Thresholds as specified within the Model Testing Guidance(The material losses are defined as follows - Secured products: [NCL$ or GCL$] > $10MM]).Backtest results are also provided for SFO portfolio. Please note the SFO portfolio(Table 3) had borrowed the RM Model for estimating the gross credit losses. Detailed justification on why a separate model was not built for SFO, can be found in Section 1.1. NCL is not a relevant metric for the SFO book. All errors for the SFO portfolio are within MRM's thresholds.

Table – 1 Twenty-seven (27) month Backtest Assessment (CMI)

Backtest Period	Metric	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold
Stress 200801	GCL(AMT)	18,905.65	17,902.63	(1,003.03)	-5.31%	25%
27 months	NCL(AMT)	5,942.94	6,146.32	203.38	3.42%	25%
Recent 201603	GCL(AMT)	1,023.53	1,066.98	43.45	4.24%	25%
27 months	NCL(AMT)	49.9	53.27	3.37	6.76%	25%

Table – 2 Twenty-seven (27) month Backtest Assessment (CPB)

Backtest Period	Metric	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold
Stress 200801	GCL(AMT)	225.82	359.58	133.76	59.23%	25%
27 months	NCL(AMT)	6.40	63.18	56.77	886.47%	25%
Recent 201603	GCL(AMT)	118.06	106.54	(-11.53)	-9.76%	25%
27 months	NCL(AMT)	5.38	6.54	1.16	21.60%	25%

Table – 3 Twenty-seven (27) month Backtest Assessment (SFO)

Backtest Period	Metric	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold
Stress 200801 27 months	GCL(AMT)	33,267.56	40,117.93	6,850.37	20.59%	25%
Recent 201603 27 months	GCL(AMT)	1,709.89	1,590.74	(119.14)	-6.97%	25%

The characteristic analyses have been executed to demonstrate the model's risk rank ordering, in tune with the model's segmentation scheme. It includes major risk factors around borrower FICO, combined loan-to-value (CLTV), loan delinquency state, region and product type (ARM vs Fixed) that potentially affect the model performance.
Sensitivity analysis reveals improved model sensitivity comparing stress and non-stress periods forecast. Since the actual losses have been trending down in recent times due to increasing HPI,

improved portfolio composition mix and declining unemployment, the model's base forecast has decreased as compared to one year ago. Further, the targeted improvements made this year have also narrowed the model errors in both stress and recent periods, which contributed to the higher sensitivity ratio. Presented here are the Sensitivity results for the CMI, CPB and SFO portfolios for the twenty-seven month horizon.

Table – 4 Twenty-seven (27) month Sensitivity Assessment (CMI)

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	592.77	1	17.23	1
Stress	1,572.53	2.65	237.91	13.8

Table – 5 Twenty-seven (27) month Sensitivity Assessment (CPB)

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	117.21	1	15.32	1
Stress	386.61	3.3	89.37	5.83

Table – 6 Twenty-seven (27) month Sensitivity Assessment (SFO)

201806 Snapshot 27 months	
Predicted GCL ($MM)	Ratio vs. Base
565.45	1
1,530.03	2.71

Model Enhancements – 2019 Method A RM Model Suite

The significant improvements made in the model's predictive accuracy for the 2019 CCAR process were driven by targeted changes that were made to the 1) Modeling Data, 2) Modeling Methodology and 3) Production Code. Presented below is the summary of all modeling enhancements made on the 2019 Method A RM model.

- The model development data has been augmented with the inclusion of the most recent time period until end of 2017 to account for the changes to the current portfolio mix which is reflective of a stronger credit profile and improving macro-economic conditions
- The segmentation analysis was conducted using quantitative factual based analytical tools (decision tree) which led to the introduction of several new interaction variables that would be able to capture the model's recent and stress periods performance better across granular segments. Please refer to Section 5.1.3 for additional details.
- A methodical approach was utilized to segregate between the modeled vs. non-modeled transitions. Because of adopting this new approach, some of the off-diagonal transitions were now modeled and additionally separate logics were introduced to systematically account for the

stress vs recent period for segments of interest (Not-in-trial vs In-trial Government vs In-trial Conventional) differentials. Please refer to Section 5.1.2 for additional details.

- A new distressed valuation haircut logic (henceforth will be called DV logic) was created which was more sensitive to macro-economic factors compared to last year's model
- The LGD modeling framework was changed from modeling losses based on their disposition types to modeling losses based on their actual occurrence (full vs partial vs. zero losses). This changed helped address the challenges associated with forecasting the future loss behavior by loss disposition types, which are more like strategic business initiatives that cannot be predicted in advance. The new framework also remediated the challenges associated with modeling loans with zero losses and improved the model's sensitivity (and monotonicity) to borrower LTV, an important measure of the existing collateral.
- Enhancements were made on the partial charge-off (PCO) logic to initialize the write-down process based on correct accounting principles and then apply the FFIEC calculation reflecting both the updated HPI and the mark-to-market LTV.
- Key code package implementation enhancements were introduced in the 2019 model project, including code package modularization, improved code verification procedures and automation efforts for report and template generation. Components of the model package can also now be run independently to allow faster testing of code changes to specific modules as long as pre-requisite code steps have been run at least once.

Model Stakeholders and Governance

The North America Mortgage Method A CCAR models were developed by the in-business Model Development Team known as Consumer Analytics Modeling Unit (CAMU). These models were developed with extensive consultation with model end users in Loss Forecasting, Portfolio Risk Management, Senior Risk Management, and Independent Risk. At the initiation of the 2019 model planning process, extensive time was spent to understand the scope of this year's model development work taking into account all internal and external feedback (and known limitations) of the prior year's model suite. This CCAR Annual Plan was vetted internally in April – May 2018 with Model Risk Management and Independent Risk who approved the proposed timeline for documentation submission and subsequent validation.

A series of internal checkpoints, coupled with presentations to model stakeholders, sponsors, and Independent Risk, occur throughout the model development lifecycle to ensure predictive credit risk models are developed in compliance with all applicable Bank policies. All CCAR models are then independently reviewed for conceptual soundness by Independent Risk and Model Risk Management respectively.

Model Development Document Sign-off

[By submitting this MDD in BankRisk MRMS for validation, the Model Sponsor acknowledges his/her sign-off on the MDD. If submission is performed by a person other than the Model Sponsor (a delegate), attach evidence of delegation from the Model Sponsor.]

Evidence attached in 0.0 Model Sponsor Delegation Signoff 092718.msg, 0.0 Model Development Document Signoff 20181121.msg, 0.0 Model Development Document Signoff 20181221.msg.

Table of Contents

1. Model Scope, Purpose and Use

1.1 Model Objectives and Business Purpose

[Model Sponsor must clearly state the model objective in terms of the below:

- Model Output (e.g. PD, Delinquency buckets, Balances, Revenues, NCL etc.)
- Exact model usage (for e.g. CCAR, CECL, ICCAP etc.),
- Forecast horizon for each of the usages (short, medium and long term) and
- Applicable scenarios (for e.g. base, stress etc.) for each of the model output and usage
- Forecast level (for e.g. account level, cohort level, portfolio level etc.)

Also, specify whether the model is intended to be a champion, challenger or benchmark model. The model usage grid should be provided in the below format.]

Table 1: Model Usage Grid

Sr. No	Model Usage	Model Outputs used	Forecast Horizon	Applicable scenarios	Forecast level

Model Sponsor is required to provide the model testing results per the model usage grid and in line with the model testing guidance.

Help the reader understand the model objectives and business purpose using non-technical language and high-level visualization tools (pictures, graphs, etc.) when possible.

Example: *This Modeling Approach is designed to estimate Net Credit Losses (NCL) in the credit portfolio <segment/ portfolio name> under baseline and stressed economic scenarios. It is used for regulatory mandated stress testing (CCAR and DFAST) as well as Bank's internal stress testing scenarios (e.g., BSST's and GSST's). The model is used to forecast losses for all five CCAR macro-economic scenarios (BHC Baseline, BHC Stress, FRB Baseline, FRB Adverse, FRB Severely Adverse), and is intended to be a champion model.*

In addition to the CCAR/DFAST usage, the model is also planned to be used for additional usages such as CECL, ICAAP etc. Please see details in the model usage grid below:

Table 2: Model Usage Grid - Example

Sr.No	Model Usage	Model Output	Forecast Horizon	Applicable scenario	Forecast level
1	CCAR	PD, GCL	Medium, 9Q	Base and Stress	Account level
2	CECL	NCL	Long, 15Q	Base	Portfolio level

3	Business Planning	NCL, Delinquency Buckets	Short, 4Q	Base only	Portfolio level
			Medium, 9Q		
4	ICAAP	NCL	Medium, 12Q	Stress only	Account level

*[*Required: Please provide response here]*

The Federal Reserve's annual Comprehensive Capital Analysis and Review (CCAR) is an intensive assessment of the capital adequacy of large, complex U.S. bank holding companies (BHCs). The financial reforms which emerged in the aftermath of the 2007-2008 financial crises triggered an evolution of regulatory policy changes which greatly affected how banks manage their capital, liquidity, leverage and other critical funding requirements. At the forefront of this massive wave of changes, stress testing emerged as a major supervisory tool for the regulators to gauge the resilience of the banking system to withstand severely adverse financial and macro-economic shocks. Given the heightened expectations, **the Federal Reserve expects the BHCs to have sufficient capital to withstand a severely adverse operating environment and continue to be able to lend to households and businesses, continue operations, maintain ready access to funding, and meet obligations of creditors and counterparties.**

Currently all Bank holding companies with consolidated assets in excess of $50 billion, as well as nonbank financial institutions designated as systemically important by the Financial Stability Oversight Council (FSOC), are all required to undergo the Comprehensive Capital Analysis and Review (CCAR) annually.

Quite similar to the CCAR Stress Testing, as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd Frank Act), the Dodd-Frank Act Stress Testing (DFAST), requires all federally regulated financial institutions with assets in excess of $10 billion to conduct company-run stress tests (FRB, 2014b, 2015b). The DFAST requirements are quite similar to CCAR and tend to serves many of the same core purposes, although these tests are applied to a much larger set of firms. While the results of the DFAST are not directly used to set capital requirements, they are critical means for supervisory oversight and enforcing tangible actions. Regulatory stress test testing has already been vetted as an effective tool to assess the robustness of the banking structure through: reducing information asymmetry among market participants and incentivizing the bank towards maintaining holistic risk-management practices.

The 2019 Method A suite of North America Mortgage models have been re-developed from its prior generation (2017) to not only address the annual CCAR/DFAST requirements but also the feedback/limitations posed on the prior models [both Champion -2017 Method A – Model IDs – 110324/110389 and Challenger – 2018 Method B – Model IDs # 128155 / 40US824] by the internal/external model reviewers' and regulators. Please note that the 2018 Method B model suite did not separate out the modified loans from the non-modified loans. Instead three models were developed at the product level (first lien mortgage, FHREL and HELOC) and each product level model included the modified loans in their scope. The LGD/loss severity component have been rolled under the PD component for each product level model. Due to the similarity of vast majority of the limitations across Method A and Method B model suites, in following sections, CAMU has segregated model limitations by their descriptive type (instead of by specific model suite) and added narrative around how each limitation has been handled in the 2019 Method A redevelopment efforts. The redeveloped model suite, which constitutes an updated Model Development Documentation (MDD), combines statistical models with academic literature and other research papers on stress testing to generate quantitative assessments on the bank's financial well-being during exceptional, but plausible future stress events.

The redeveloped model suite, which constitutes an updated Model Development Documentation Template (MDD), combines statistical models with academic literature and other research papers on stress testing to generate quantitative assessments on the bank's financial well-being during exceptional, but plausible future stress events.

This redeveloped Method A RM model suite is intended to be used for the annual CCAR/DFAST cycle for all FRB and BHC[BHC Baseline, BHC Stress, FRB Baseline, FRB Adverse, FRB Severely Adverse] scenarios, mid-year CCAR cycle BHC scenarios for all loss forecasting exercises that include a) GCL (Gross Credit Loss) - dollar value and as a percent of ending net receivables (ENR), b) NCL (Net Credit Losses) – dollar value and as a percent of ending net receivables (ENR) and c) month over month delinquency projections for the nine (9) quarters of the forecast window.

Apart from meeting the CCAR specific mandates, the model can also be used for Non-CCAR initiatives which include the Allowances for Loan for Lease Losses and Loan loss reserve processes (ALLL and LLR). For ALLL and LLR uses, loss forecasts are generated by running the model in a flat macroeconomic scenario. A flat macroeconomic scenario is created by retaining fixed values for macroeconomic parameters as of snapshot date. The model is also used for Business Planning purposes to develop a business direction (sales, fulfillment, servicing) to achieve the desired earning or return target, and the business decision is partially based on the NCL/Delq/LLR forecast. Further, the model can also be used for managing the risk appetite of the mortgage business as set forth in Bank's Global Consumer Credit Fraud Risk Policy (GCCFRP) that in turn can help maintain a prudent capital allocation process through process improvements initiatives or exploring new venues of growth. The Risk Appetite Framework (RAF) is a core component of Global Consumer Bank's Risk Governance Framework (RGF) which is part of Bank's enterprise wide approach to managing all risk types including Credit, Price/Market, Interest rate, Operational, Strategic, Reputational, Liquidity and Compliance Risk. The purpose of RAFs is to clearly articulate and identify the key risk that arise from business strategy and activities, discuss boundaries around all key risks, including behavior expectations for qualitative risks, and metrics, limits or thresholds for monitoring and managing quantitative risks, and ensuring the right policies, governance and control processes, and risk limits are in place, and operate effectively, to manage those risks. All pertinent details around Bank's Risk Appetite Framework (RAF) are provided in the Global Consumer Credit Fraud Risk Policy (GCCFRP) annexure made available in the Appendix section.

In summary, this model is proposed to be utilized for both CCAR and non-CCAR business purposes. A full list of proposed model uses is provided in Model Usage Grid (see table 1.1.1 Model Usage Grid below).

CAMU will work with model user teams to provide additional details as required for the listed uses during documentation of model overlays. Model overlay documents are independently reviewed by MRM / IRMO prior to using the model.

Table 1.1.1 Model Usage Grid

No.	Model Usage	Model Output	Forecast Horizon	Applicable scenario(s)	Forecast Level	CMI RM	CPB RM	SFO RM
1	**Annual CCAR Capital Adequacy Assessment (CCAR) and Mid Cycle Stress testing (MCST)**	1. Delinquency Buckets \$ 2. Net Credit Loss (NCL)	Medium	Base and Stress	Portfolio	Applicable	Applicable	Applicable
2	**Repurchase Risk[2]**	1. Gross Credit Loss[1] (GCL) \$	Medium, Long	Base, Stress and Flat	Portfolio	Applicable	Applicable	Applicable
3	**Current Expected Credit Loss (CECL) for:** - **Modified (MOD) loans – FAS114** - **Never-modified loans – FAS5**	1. GCL units 2. Voluntary Prepayment (VP) Units 3. NCL	Long	Base	Portfolio	Applicable	Applicable	Applicable
4	**Allowance for Loan and Lease Losses (ALLL) and Loan Loss Reserve (LLR)**	1. GCL units 2. VP units 3. NCL	Medium, Long	Flat	Account and Portfolio	Applicable	Applicable	Applicable
5	**Other BAU purposes including:** - **Business Planning (including SFO)** - **Selected assets disposition for asset sale** - **Collection strategy and action** - **Capital Markets Planning** - **Risk Appetite (including New Loan Performance monitoring)**	1. Delinquency Buckets units 2. Delinquency buckets \$ 3. GCL units 4. GCL \$ 5. GCL%ENR 6. VP units 7. VP \$ 8. NCL 9. NCL%ENR	Short, Medium, Long	Base and Stress	Account and Portfolio	Applicable	Applicable	Applicable

[1] Gross Credit Loss is calculated as first time 180+ days past due (DPD) and direct IVP (IVP from lower than 180DPD) for never 180DPD loans, and IVP for ever-180DPD loans

[2] For additional details on the exact usage pertaining to Repurchase Risk Assessment, please refer to REL Repurchase Model ID 111427 MDD in iMRMS

To provide some historical context, it is important to note here that while the 2017 Method A PD model suite was approved as the Champion model and determined to be both functionally sound and technically robust for use in CCAR forecasting for the NA Mortgage portfolio; the model tended to display conservatism for recent period backtests over nine (9) quarters. This performance exception was in contrast with the model ability to display strong back test performance during stress as well as in the post –stress periods.

Given the feedback received from the federal regulators and independent model reviewers' on the 2017 model interjected with the model's otherwise robust performance, the 2018 CCAR efforts focused on refining the model's backtest process affected by the action loans and developing a deeper understanding on the model's transition movements in a multi-period forecast, instead of doing a complete model redevelopment. All of these afore-mentioned enhancements to backtesting logic, delinquency code improvement and granular level HPI update were submitted as part of a Change Addendum for the 2018 CCAR submission that was independently validated and approved.

The Model Validation Report (MVRT) provided at the end of 2018 submission process on the Model Change Addendum (CAs) and the Model Development Document (MDD) for the Method A Method B first Mortgage PD and Severity models, cited technical and functional limitations and corresponding recommendations that pertained to enhancing the model segmentation schema across more granular business portfolios, products and risk drivers, enhancing the partial charge-off logic, standardize the process around the modeling of transition vs non-transition equations with appropriate supportive empirical evidence(S) across granular population cohorts and take targeted approaches to mitigate some of the model's shortcomings' around recent period over estimation, with incorporation of additional risk drivers pertaining to macroeconomic factors and LTV. These limitations and recommendations were thoroughly evaluated during the re-development effort in preparation for the 2019 CCAR submission.

Enhancements introduced in model

Presented below is a brief synopsis of the itemized improvements that were introduced as part of the 2019 Method A RM model suite -

1. **Data Alignment across business entities and products**

One of the key improvements implemented as part of the 2019 CCAR cycle was the adoption of clear and consistent product definitions across both the Champion (Method A) and Challenger (Method C) model suites. CAMU made an effort this modeling cycle to align business and product characterizations within the model across all model usages pertaining to CCAR reporting, non-CCAR BAU usages. The development of the product definition code logic itself was a collaborative effort between CAMU, REL DataMart team and the model's business users. The creation of one uniform product identification logic has been deemed to facilitate the comparison across independent model suites as part of model benchmarking processes (Benchmarking is due end of 2019Q1. Please refer to Section 6.5 for additional details).

2. Incorporation of recent period data in the model development pool

As per limitation # 19546, it was recommended to include recent period data as part of model development for both the PD and severity models. Given this cited limitation, both the PD and LGD models extended their development samples to include the most recent timeframes. While there was also a recommendation made to include the pre-2008 data as part of severity model development sample, a deep dive analysis in collaboration with the REL DataMart team revealed discrepancies within the Risk Loss data prior to 2008, attributable to either manual entry errors or top line adjustments to actual loss numbers. Since the reconciliation error between Risk DataMart's Loss file and PEARL (master data file) for the pre-2008 loss data exceeded well over the 1.5% threshold set by REL DataMart team, it was considered prudent to not use the pre-2008 data as part of the Severity model development. For additional details, please look at the attachment – '1.1 Model development data sources and Mortgage Transformation considerations' for pertinent details. The Method A RM PD model data extended from Feb-2006 until Dec- 2017 while the LGD model's development data extended from Jan 2008 until June 2017. The inclusion of the recent data points helped with capturing the recent portfolio trends and origination profile(s). Additional details on the data trends are discussed in Sections 4.4 and 4.5 of the MDD.

3. Quantitative, factual analyses that support the modeled vs. non-modeled transitions

As per limitation # 19638, 19775, it was cited that the 2017 first lien PD model assumptions on non-modeled transitions were counter-cyclical. Given this limitation, the model development team developed a systematic approach based on historical data movements, key risk driver movements and macro-economic trends to empirically rationalize the intuition behind modeled and non-modeled transitions. To add context, as part of the 2019 PD model redevelopments all transitions were evaluated to distinguish and separate out the modeled vs modeled transitions using a statistical waterfall approach that combined the transition volume, contribution volume between transitions, volatility associated with each transition along with their measured C-statistics to deem their eligibility for being modeled as equations . For a transition that was deemed as non-modeled, an empirical lookup logic was created that accounted for the base vs stress, in-trial vs not-in trial and government vs conventional differentials to account for their differentiated performances across the afore-mentioned parameters. Further this lookup logic was tested using a 10% shock to illustrate the model sensitivity to these assumptions. This approach had two-fold advantages-

a) It was able to capture the differences in the model's performance across stress and recent periods with the increase in model's sensitivity across both these individual timeframes,
b) Balanced the model's fit with the limited data availability

For additional details on the modeled vs. non-modeled transitions, please refer to Section 5.1.2 of the MDD as well as supporting attachments. All these enhancements helped remediate the model's stress and recent period trigger breaches (limitations # 16735, 16736).

4. Improvements to PCO logic

One of the areas that were scrutinized intensively as part of the 2018 CCAR review process, both as part of the Method A Code Review and the model's functional validation (Limitation # 19736) was the Partial Charge-off (PCO) logic. It was noted by independent model reviewers that the PCO logic charged to loans was based on the decline in prices, without any conceding reference to the loan balances or LTV which in turn would tend to overstate losses especially for the low CLTV populations. These comments/reviews led CAMU, in collaboration with the relevant business partners, to obtain details on accounting loss recognition and write-down practices. Based on the information shared, CAMU initialized the PCO write-down using historical data, applied the relevant policy (FFIEC) calculation based on the principal balance and BPO estimate reflecting not only the updated HPI but also the mark-to-market LTV of a given loan and capped the FFIEC charge at the loan balance to preserve the monotonicity of the write-down process. For additional details on this modeling enhancement and the related impact on the PD model results, please refer to the attachment '1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY'.

5. Refined Segmentation Analysis

Based on the feedback from independent model reviewers as well as internal stakeholders regarding limitation #19775, the 2019 model segmentation was based on a factual, quantitative decision tree analyses. The decision tree analyses helped to confirm the model's existing segmentation and identify new segments of interest based on their differentiated sensitivities to key risk drivers. Appropriate quantitative evidence and justification were provided for the segments that were combined together as part of coping strategy. Once the main segments were identified, additional analyses were conducted in terms of bivariate and cluster analyses to check the sensitivities of these segments across key risk driver such as MTM LTV, FICO, HPI growth, etc. Subsequently given the observed sensitivities, enhancements were incorporated either in terms of new modeled equations and/or inclusion of interaction terms, to improve the model's performance. Additional details on the segmentation analysis executed in preparation for the upcoming 2019 CCAR submission is presented in Section 5.3 of this document.

6. Improved model's sensitivity to recent and stress periods

As per limitations # 19638,16735, 16736 it was cited that both the 2017 first lien PD and 2018 first Mortgage modes were not sufficiently sensitive to stress and recent periods. To remediate these concerns, the model development team made several enhancements, all of which are listed as below-

a) Introduced and tested new macro-economic variables (S&P500 and VIX) to capture the model's sensitivity to the equity market performance;
b) Deep-dive analysis was conducted and variable transformations (change of change, interaction) were introduced to capture the model's stress peak losses and improve the model's VP equation sensitivity to macro-economic factors and to capture the model's recent period BUK7 -> IVP increase.
c) Introduced new HPI burn-in and VP burn-out attributes to capture the portfolio composition changes in response to the changes in the broader macro-economic environment.

All of these enhancements have been explained in length in Section 5.1.2 of the MDD.

7. Replaced the 2016 DVM model with a simplified DV haircut lookup logic (DV Logic)

In response to the functional validation limitation # 19516 imposed on the existing Method A DV model, the 2019 CCAR process replaced the existing model with a haircut logic that was customized based on the segments modeled (PDV) across business cycles (stress vs non-stress) and key risk indicators. The DV model that was in existence in prior years, although produced robust results, was deemed to have had an overly simplistic variable selection process that had no business or loan characteristic drivers other than home price index (HPI). Further as part of the Bank's ongoing strategic Mortgage Transformation initiative (please see attachment – '1.1 Model development data sources and Mortgage Transformation considerations for more details) which posed challenges in terms of accessing some historic data files, it was considered prudent to introduce a haircut that would utilize a monthly refreshed property valuation table that would not only be based on more accurate valuations but would also be consistent across the model's stated business usages. This new haircut logic has been tested for all loan level specific attributes and across stress and non-stress periods to ensure model's appropriate responsiveness for all modeled segments. All details pertaining to the DV logic have been discussed in Section 5.1.2 of the MDD.

8. LGD Modeling Approach for first lien loans

The LGD modeling methodology for the first lien loans was redeveloped to align with the methodology used for second liens. While there were no outstanding limitation to prompt this change, this modification in the modeling approach was considered necessary given the evolution in the trend of the loss disposition types in recent time periods. To get additional insights justifying the change in the modeling approach, please refer to Section 3.1 of the MDD. The new framework for the first lien loans would leverage a two stage approach towards modeling losses wherein the first stage would segment the losses based on their outcome [full, partial and zero losses] and the second stage would model the loss severity for the partial losses only. This proposed framework will help mitigate the shortcomings associated with forecasting the future loss behavior by loss disposition types which are more like a strategic business initiatives that cannot be predicted in advance and also remediated the challenges associated with modeling loans with zero losses. Additional details on the new LGD framework in presented in Chapter 5 of this MDD.

9. LGD modeling approach enhancement for VA loans/Loans w Zero losses

All VA loans are included in the model forecasting (limitation # 14918) whereby the loan losses are estimated based on the Method A RM model and then adjusted by the appropriate guarantee amount based on the appropriate loan sizes. The guaranteed amounts are based on the relevant accounting policy as supplied by the business. The separate process that has been adopted for the VA loans helps to remedy the limitation # 19546 that suggested seeking a different specification for the VA loans. Given the change in the modeling approach, all zero loss loans are included in the development pool and used as part of model estimation process. The inclusion of zero loss loans remediates limitation # 19546 issued on the prior model suite that tended to exclude all zero loss loans from model development process. For additional details on the justification to include zero loss loans, please refer to Section 3.1 of the MDD.

10. Enhanced production code

Based on CAMU self-identified process improvements review for the 2018 Method A production code the below enhancements were completed for the proposed 2019 Method A production code:

- <u>Code package modularization</u>
  Several code sets in the model package have been modularized to enable separation of model customization settings from the model engine (loan scoring and calculation codes). Examples of customization settings include type of model run (back test / sensitivity), snapshot date, output path, ON/OFF selection of stress setting for model parameters, and selection of pre-defined sensitivity test scenarios to be executed. Components of the model package can also now be run independently to allow faster testing of code changes to specific modules (as long as pre-requisite code steps have been run at least once). Code package modularization allows easier implementation of End User Computing (EUC) controls since the centralized model engine code sets are now independent of customization codes.

- <u>Alignment of reporting indicators and reporting process automation</u>
  The 2019 model package includes a standardized list of loan indicators to allow consistent aggregation, tracking and benchmarking of model performance across segments of interest along with a complete attribution of projected end-to-end losses (limitation # 19545). These indicators facilitate discussions with Business teams and will enable improved benchmarking processes with the alternative Method C model suite (Challenger model). Extensive automation initiatives were introduced to enable completion of MRM templates for back testing and sensitivity analysis to accompany model document submission. Automation efforts included creation of sas codes and vba macros to allow consistent and repeatable preparation of reports from model outputs.

- <u>Code package verification process</u>
  Code implementation and code execution for report generation were carried out by separate sub-teams within CAMU's production team, along with a third step of independent verification of final reports prepared from model outputs(limitations # 16777, 16782, 16582). This approach allowed multiple rounds of verification of the code package and enabled the team to identify updates needed to prepare the code package for delivery to independent validation (MRM) and business user teams.

- New output fields for Gross Credit Loss
  Prior generations of the Method A model package did not contain a direct model output field for Gross Credit Loss (GCL), and model users were required to use Involuntary Payoff (IVP) and Exposure at Default (EAD) metrics to compute GCL measures. GCL unit and dollar amount metrics are now included in the list of model output fields.

11. Model Usage Grid

In collaboration with the business users to holistically assimilate all model usages across all business entities and products, model developers spent a considerable amount of time, generating a set of consistent reportable metrics that can be readily used across all dimensions. The appropriateness of all modeled outputs (CCAR & non-CCAR) were assessed for given model usages for each specific modeled scenario (base vs stress), across specific forecasting horizon(s) and level (portfolio vs account level) as stipulated within the Model Testing Guidance (MRM). Please note that the model usage grid has been completed based on the collaborative discussions with the model end users around the purported model usages and to ensure compliance with the model usages requirements, set forth within the June 2018 release of the Model Testing Guidance report. For a complete overview of the 2019 Method A RM Model Usages, please refer to table 1.1.1 Model Usage Grid. Furthermore a complete benchmarking analysis for each model usage (in response to limitation # 16609) would be conducted once MRM completes the eligibility assessment for the alternate Method C framework and results from this would be submitted by May-2019, as decided in a mutual agreement between MRM and CAMU as documented in the Annual Plan.

An important segway worthy of discussion here is regarding the applicability of this model suite towards estimating losses for BankMortgage's Serviced for Other (SFO) portfolio. To provide some context, Bank's SFO portfolio reflects the sold and serviced loans that Bank retains Mortgage Servicing Rights (MSR) which are considered off balance sheet. The Serviced portfolio consists of all loans that Bank services which were acquired through secondary market trading. Loan servicers are normally compensated by receiving a percentage of the unpaid balance on the loans they service. The fee rate can be anywhere from one to forty-four basis points depending on the size of the loan, and the level of service required. Those services can include (but are not limited to) statements, impounds, collections, tax reporting, and other requirements.

As of 2017Q2, BankMortgage made the decision to exit the mortgage servicing business in a gradual phased out manner (please refer to attachment '1.1 Model development data sources and Mortgage Transformation considerations'). The decision was based on a forward looking view to create a leaner organization with a focus towards integrating retail banking operations with mortgage originations, coupled with enhanced digital capabilities that would better position the business for the future and organic growth. Given this strategic business decision, the size of the SFO portfolio has rapidly decreased over the last fifteen months.

During this model usage assessment process, a considerable amount of time was spent in evaluating the need/rationale(s) to build a separate model for the SFO portfolio. In all conceding dialogues with the

SFO Risk Management, it was mutually decided NOT to build a separate model for the SFO portfolio, which is consistent with how losses have been forecasted in recent years. The Model Sponsor proposes that the SFO portfolio will continue to borrow the Method A model suite for all its reporting and forecasting needs. This decision was strategized based on the following rationales-

1. Loans originated and the decision to sell or retain on the balance sheet is a business decision that changes over time based on portfolio mix, concentration, and risk appetite. CAMU and BankMortgage Risk Leadership have discussed that it will not be feasible for a model to continuously incorporate these ongoing changes based on evolving business strategies. The servicing strategy has included in recent years' loans sales of non-performing, MSR exposures, and even exiting the servicing business altogether. Bank achieved a large asset sale in 2017, which dramatically reduced the SFO portfolio. BankMortgage Risk Leadership and Model Sponsors believe that a borrowed model including historical portfolio performance best suits the needs for forward looking projections which will continue to yield conservatism.

2. A separate SFO model does not meet the business's Cost vs. Benefit justifications based on Bank exposure to only the MSR value of SFO loans ($0.57B fair value, as of April-2018). Accurate model for the Risk book separately takes a priority over sold-serviced SFO book, noting the strategic Mortgage Transformation initiative that Bank undertook Q1-2017. Additional details on the Mortgage Transformation initiative can be found in attachment '1.1 Model development data sources and Mortgage Transformation considerations'. With the resource constraint imposed by the ongoing Mortgage Transformation initiative, and based on CAMU's annual plan prioritization, CAMU, in conjunction with the business users decided not to develop a separate model for the SFO portfolio.

3. The SFO book typically consists of loans, which are relegated to the sale status due to some material defects arising from either manufacturing defects or non-compliance with investor standards. As investor standards may differ from Bank's global credit risk policies, it was considered a prudent modeling approach to not combine the SFO and Risk portfolios together.

It is important to note here that while a borrowed model approach has been utilized for the SFO performance reporting, CAMU worked in parallel with the model users to address all future borrowed model performance issues noting a very conservative forecast. All pertinent details surrounding the SFO model performance can be found in Chapter 6 of the MDD.

Compliance with CECL (Current Expected Credit Losses)

Under the new CECL requirements, financial institutions are required to use historical information, current conditions and reasonable forecasts to estimate the expected loss over the life of the loan. While CECL represents a significant change in accounting for the allowance of loan loss/leases reserves (ALLL) calculations from the prior FAS5/FAS114 accounting policies, the current credit risk models as built by the CAMU team for the annual/semi-annual stress testing (CCAR/DFAST) of the US real estate portfolios already provides some elements that can be potentially leveraged for CECL. To be more specific, the underlying transition matrix, the loss severity component and expected loss (EL) framework estimation methodologies, among others, are all built at a loan-level that can be holistically assimilated to estimate the losses over the life of a loan to fulfil CECL requirements. Some key areas of convergence between the current modeling methodology and its overlap with the CECL requirements are highlighted as below –

Modeling Components	Methodology deployed for fulfilling CCAR, DFAST/Non-CCAR usages	CECL Requirements	Overlap
Probability of Default model	Point-in-time PD based on the month-over-month transition framework; Monthly PDs can be aggregated using the Expected Probability Approach to get the lifetime PD	Life-of-loan PD; Point-in-time PD	Yes
Loss Given Default(LGD)	Life-of-loan	Life-of-loan	Yes
Other Factors(Macroeconomic)	Forecasts have to done on 13 quarters to meet CCAR/DFAST requirements. Further forecasts are available over the life of the loan based on MRM's stipulated forecasting horizons	Macroeconomic forecasts need to align with the same time horizons used in assessing PD and LGD; assumptions. Further relationships to the portfolio composition changes and the allowances needed are to be demonstrated	Yes, all macro-economic forecasts align with the time horizons. Further, the model users review the model outputs in conjunction with the evolving economy and internal/external policy changes and apply overlays to modeled outputs as deemed necessary

As noted above the current model is capable of being applied for CECL forecasting. CAMU is in discussions with model user teams from the Loss Forecasting groups to finalize scenario requirements for CECL uses of the model. Model uses specified in Table 1.1.1 Model Usage Grid include CAMU's proposal to use this model for CECL forecasting. CAMU will work with model user teams to provide additional details specific to CECL forecasting uses during documentation of model overlays. Model overlay documents are independently reviewed by MRM / IRMO prior to using the model.

12. Model Benchmarking

As per limitations #s (16578,16737, 16755, 16579, 16580, 16581, 14919, 14921, 16737, 16755) the Method A RM model suite is a completely re-developed model suite which would be benchmarked against a completely new Method C (Challenger) model suite. The Method C model suite (due for submission January 2019) uses a Hazard modeling framework which is substantially different from the transition model framework. The Hazard framework models the termination events in a competing risk framework. In other words, for each month of a scenario forecast, the model's competing risk survival methodology predicts the probabilities of the loan termination events – ICP, VP first time BUK7 and loans which are either sold & released or sold & serviced, conditional on the survival of the account to date. Based on the submission of the Annual Review Plan to MRM and the corresponding communication, it was mutually agreed that the complete benchmarking analysis (for all stated model usages) would be due for submission, after the Method C model suite has been validated by the model technical + functional stakeholders. For additional details, please refer to the attachment '1.1 RE Meeting Minutes CAMU modeling plan and limitation compensating control discussion'. The 2019 RM model has also been benchmarked to its prior version (2017 Method A 1st Lien model) to get an accurate assessment of uplift in the model performance, given this year's targeted enhancement efforts. A complete comparison of this current model with the prior version has been provided in Section 6.5 of the MDD.

13. Granular Level Analytics

In order to satisfy limitation #16581, CAMU has created comprehensive characteristic analyses based on the relevant segments of interest, as confirmed and validated by the model's segmentation scheme. Additionally, prudent care would be taken to ensure that the all analyses including the end-to-end results of the complete PD + EAD + LGD model suite and outputs (including metrics definition and their intuitive appeal) conform to the requirements set forth within the Model Testing Guidance Report. All backtesting and sensitivity analyses would adhere to MRM's Model Testing Guidance requirements on the desired level of reporting granularity and precision.

14. Appropriateness of Coping Strategy

In order to address limitation #s (16581, 14920, 16776) surrounding appropriateness of the coping strategies used, the modeling team conducted several deep-dive analyses to justify the approaches utilized. For the RM PD model, a decision tree analysis was leveraged to gain insights on the importance of CMI vs CPB split. Due to lack of significant performance differential between these two business entities as depicted within the tree, while it was considered prudent to combine the CMI portfolio with the CPB portfolio, the modeling team incorporated the CPB specific balance effects within the model to holistically model the CPB performance. Similarly, for the LGD model, given the sparse loss count, the model development team leveraged the CMI portfolio to derive an empirical multiplier based on the ratio of CMI base to stress losses. This multiplier was then used to calculate the add-on losses, on top of the base losses for the CPB portfolio. Additional details on this can be found in Section 5.1.2 of the MDD Further, appropriate sensitivity testing has been conducted for the Severity model to assess the impact on model output to changes in underlying macro- variables. Please note the proposed LGD model does not have any underlying assumptions.

15. Other Enhancements based on Prior Regulatory Feedback

Based on the March/April 2018 OCC-DFAST examination/feedback of the 2018 REL CCAR models, three key areas were scrutinized and recommended for next stage model development considerations. These pertained to assessing the impact from an increase in LTV (loan-to-value) requirements for the IO (interest-only) sub-group of loans, tracking the influence of equity market movements on real estate performance and assessing the potential impact of the tax reform passed December 2017 .For specific details on the questions asked and the responses provided, please refer to attachment titled '1.1 Retail_NA REL_180416_1951350_Bankgroup_Regulatory_Response' file.

Given these specific recommendations', CAMU incorporated the following changes –

1. LTV/CLTV effects within the model are used as independent attributes or interaction terms to capture the incremental credit risk of the loan. As of the July 2018 snapshot, the entire RM portfolio had 19,209 of 170,274 IO loans of where 2,047(~11%) had LTV>75%. Since the policy came into effect in early 2018, it would only impact ~ 184 loans originated in 2018. Regardless of the policy impact, CAMU carried out specific analysis during this modeling cycle, to assess the incremental risk from IO only loans. In order to understand the historical behavior of IO loans, several payment shock variables were introduced into the model specifications to capture the effect of payments shocks on borrowers' credit worthiness and overall model performance.. As such, CAMU believes that all IO specific attributes that have been introduced as part of the model specifications should be able to capture all risks associated with this sub-group of loans, especially for the IO sub- segment(s) with higher LTVs.

Table of IO by LTV_GT75

		LTV<=75%	LTV>75%	Total
NON-IO	**Frequency**	131447	19618	151065
IO	**Frequency**	17162	2047	19209
Total	**Frequency**	148609	21665	170274

2. The S&P500 index attribute was introduced in the entire Method A model suite to holistically capture the effect of equity market performance on the credit worthiness of a loan. The S&P500 index appears as a significant attribute in some key transitions. Please refer to Section 5.1.4 in the MDD for details around the finalized variables and their accompanying justification
3. Third, Internal Audit and Regulators inquired during the 2018 CCAR examination about potential impact from US Tax Reform implemented in late 2017 on Bank's loss forecasting models and modeled portfolios. In order to assess the impact of US Tax Reform on US portfolios, the businesses reviewed trends in disposable income pre

and post this change. As portrayed in the table below, no material change has been observed in disposable incomes. The changes observed, merely reflect seasonality in the data. For example, the real personal disposable income increased by 0.4% quarter-over-quarter in Apr'2018, flat compared to the change observed in Apr'2017. Additionally, it increased by 0.9% in Jan'2018, and again flat compared to the change observed in Jan'2017. Therefore, while the US businesses will continue to monitor these trends, no adjustments or overlays are being recommended at this time based on the disposable incomes.

Time period	Real Disposable Personal Income: Per Capita	% Increase
2016-07-01	41958	0.2%
2016-10-01	42162	0.5%
2017-01-01	42556	0.9%
2017-04-01	42715	0.4%
2017-07-01	42866	0.4%
2017-10-01	43027	0.4%
2018-01-01	43430	0.9%
2018-04-01	43624	0.4%

For US Real Estate Lending (REL) portfolios (covering Residential Mortgages and Home Equity lending products), disposable income effects encouraging home ownership among lower income borrowers are expected to be counteracted by effects from rising mortgage rates and tax-incentives on renting. Furthermore, the cap and reduction of mortgage/home equity interest and property taxes deduction will negatively impact housing values, especially for markets/properties which have higher state and local tax (SALT) deductions. Our assessment is that the overall impact of tax reform is marginal on US REL portfolios considering that the tax savings netted for other regions, segments, and/or income brackets in a consistently improving macro-economic environment would counter some of the adverse effects of the tax reform on certain segments and regions. Such impacts to the US REL portfolios are captured through Home Price Index (HPI) data, which is a key input in the REL loss forecasting models.

1.2 Business Scope of the Model

Which business line does the model describe?

The Method A RM model includes all mortgage loans for which BankMortgage (CMI) and Bank Private Bank (CPB) respective entities, hold the risk. A residential mortgage loan identifies any loan primarily for personal, family, or household use that is secured by a mortgage, deed of trust, or other equivalent consensual security interest on a dwelling or residential real estate. While the BankMortgage (CMI) portfolio has a more seasoned mix of loans including both pre-crisis and newer originations, Bank Private Bank (CPB) portfolio caters to a more niche, affluent population with higher revolving balances.

Irrespective of the noted differences in credit policy mix, both the afore-mentioned portfolios tend to exhibit similar key risk drivers and hence are modeled using the same Probability of Default (PD) transition framework. Additional details on transition framework are recapitulated in Chapter 5 of this document. It is important to note here that although both CPB and CMI portfolios are modeled using the same transition framework, the idiosyncratic differences between these two portfolios are captured well through the model segmentation and/or the addition of appropriate intercept/interaction effects (if a coping strategy has been used) which has been detailed in Section 5.1.3 of this document. The loss severity framework tends to differ between CMI and CPB portfolios. While a two-stage loss modeling framework has been used for the CMI residential mortgage portfolio (all liens included), an empirical lookup logic has been created for the CPB portfolio which is characterized with extremely low, yet volatile losses. This is not entirely surprising as the CPB clientele are not sensitive to the changes in business cycles and often have enough capital cushion to ride over a trough without missing any payments.

- Portfolio description

  [Please provide a detailed description of the portfolio, including size, product description, historical trends, and future/planned strategies. Please provide charts and tables wherever necessary.]

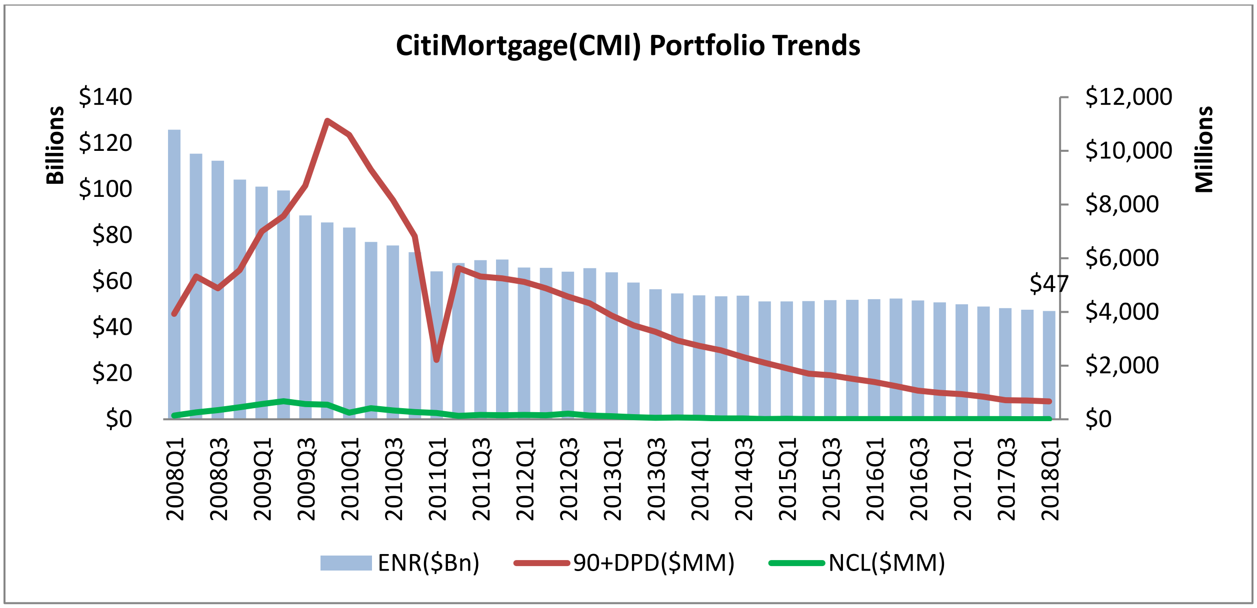
The total size for the North America RM portfolio is currently approximated at $47.1Bn in ENR for BankMortgage (CMI) and $25.6Bn in ENR for Bank Private Bank (CPB), as of the first quarter of 2018. The SFO portfolio had approximately $58Bn in volume as of Q1, 2018. Total SFO origination volume was $0.38Bn for April'18. For additional details on SFO portfolio trends, please refer to attachment '1.2 Sold_and_Serviced_Review_Jun18.ppt'.

It is important to note that that the CMI portfolio has decreased considerably over time from $125Bn (2008Q1) in managed assets at the onset of the crisis period to almost one-third of its former volume in current period. The portfolio delinquencies (90+DPD) and the credit losses (NCL) correspondingly have also gone down considerably from $3.9Bn and $134MM from the time of the crisis [2008Q1] to $0.67Bn and $6.5MM dollars respectively as of 2018'Q1. The notable drop in portfolio size and corresponding delinquencies & losses was strongly driven by the management's desire to improve portfolio mix with an onward shift towards more retail originations compared to correspondent/broker based originations, with a focus on an incentivized relationship pricing model, stringent underwriting criteria and enhanced digital capabilities. For loans originated by Correspondent Lenders for which Bank serves as the servicer, these third party originations have remained flat post financial crisis. Given the directional approach taken by Bank's management post-financial crisis, as illustrated in the prior section, the CMI portfolio got broadly segmented into two major asset classes – Corp and Holdings, based on the long-term asset class mix. The Corp portfolio is reflective of Bank's proactive risk management efforts to maximize portfolio returns encompassing a risk-adjusted growth philosophy. The core philosophy underlying the Corp portfolio growth model and strategy caters to three main objectives-

- Deepen and acquire client relationships with the Retail Bank in Bank's core markets
- Fulfil Bank's commitment to community lending to low and moderate income borrowers
- Create opportunities for saleable mortgage loans through both Direct and Distributed channels

The Holdings portfolio, on the other hand, is the de facto carrier of non-performing, riskier (sub-prime) originations which are tactfully managed for their incremental riskiness and by all means are positioned for gradual future asset sales, depending on market timing and competition.

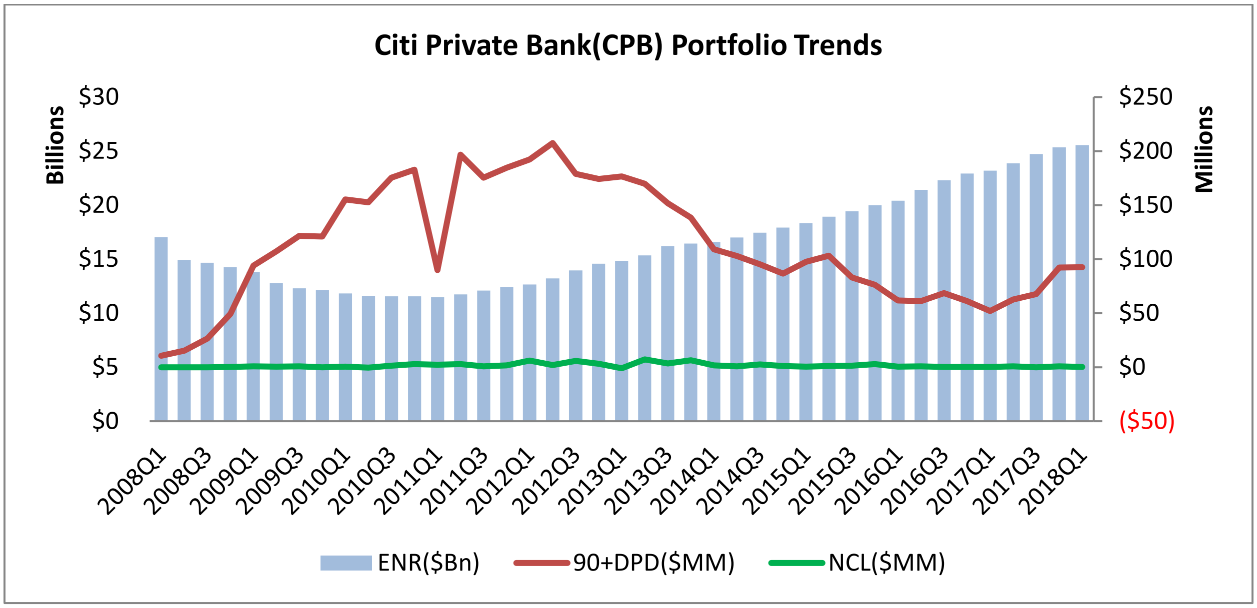
Figure 1.2.1a: CMI Portfolio Trends (2008Q1 – 2018Q1)



2011 Q1 Dip reflects Penny Mac sales

The CPB business has grown gradually from $17Bn in 2008Q1 to $25Bn as of 2018Q1 and will continue to grow in the near future, given the strategic approach to cater to private bankers (Law Firm Group, Ultra Net Worth, High Net Worth) with home financing needs for primary residences, vacation homes and investment properties, through a wide range of loans products, structure and unique ownership entities. The portfolio has witnessed sparse but volatile net losses (NCL) over time that is reflective of affluent customer bases that are usually not sensitive to business cycle movements. The overarching mission for the CPB Residential Mortgage Solutions team is to be fast and competitive in offering differentiated client solutions with a focus on creating liquidity from the real estate holdings to meet the clients global investment needs. The forward looking three year originations forecast for the CPB real estate business is also included in section 3.2 of this document.

Figure 1.2.1b: CPB Portfolio Trends (2008Q1 – 2018Q1)



2011 Q1 Dip reflects Penny Mac sales

Given Bank's gradual phase –out from the Servicing business it was deprioritized from the consideration of having a separate model and instead borrowed the Risk Book RM model for its loss assessment projections for fulfilling the regulatory and internal reporting needs, adjusted with appropriate overlays by the business users, as deemed necessary. Additional information on the SFO portfolio modeling justification can be found in Section 1.1 corresponding to the Model Usage Grid of the MDD.

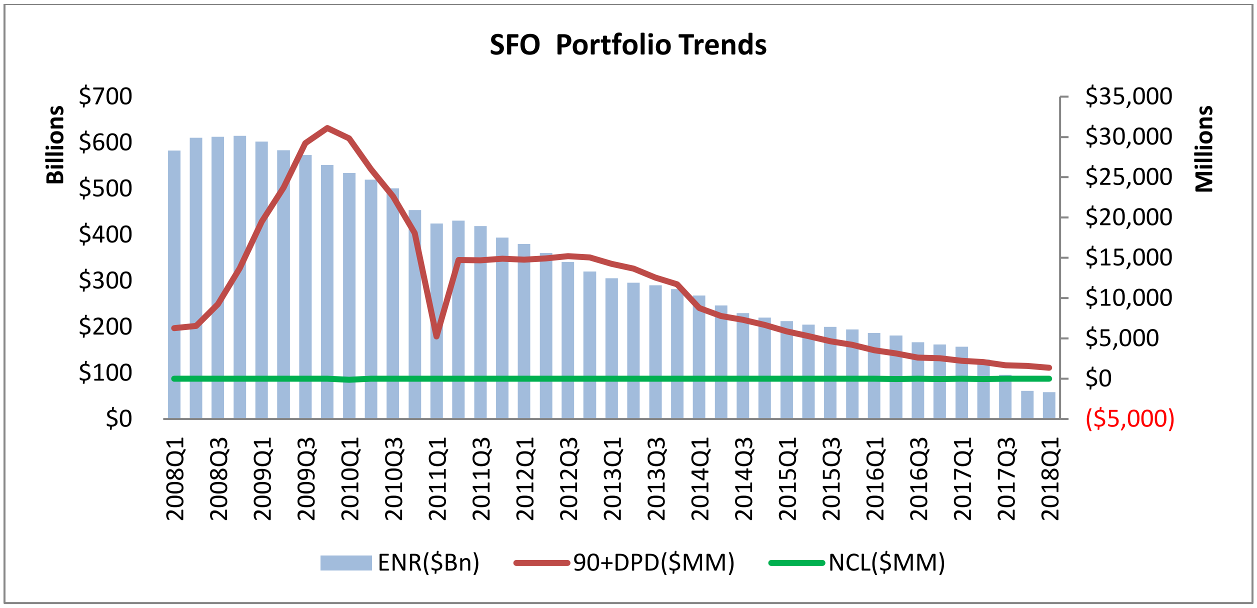
This diminishing portfolio trend has been evidenced in the chart below. Currently the SFO portfolio has approximately $58.3Bn in unpaid principal balances. Total SFO origination volume was $0.37B for April'18; of which the following reflects the net distribution-

- Non-HAR: $368MM (99% of SFO volume);
- HAR: $4MM, remains near all-time low
- Government: $0.009B (2%); 80% Purchase volume

Bank's Management continues to monitor the existing SFO portfolio in terms of its risk appetite and transfer assets, based on market timing and operational capacity.

All of these portfolios have been detailed in terms of their specific trends in the following trend charts and also within the Gating Principles in Section 3.2 of this document. All specific portfolio characteristics have been detailed by product type, volume, number of active accounts, delinquencies in order to meet Bank's required Model Risk Management (MRM) model testing guidance specifications and requirements.

Figure 1.2.1c: SFO Portfolio Trends (2008Q1 – 2018Q1)



2011 Q1 Dip reflects Penny Mac sales

All of these portfolios have been detailed in terms of their specific trends in the following trend charts and also within the Gating Principles in Section 3.2 of this document. All specific portfolio characteristics have been detailed by product type, volume, number of active accounts, delinquencies in order to meet Bank's required Model Risk Management (MRM) model testing guidance specifications and requirements.

The Method A model suite is conceptually designed to generate end-to-end loss forecasts for the North America real estate portfolios using actual macro-factors, in one comprehensive modeling unit. It is important to note that human behavior is not always rational nor are the market movements perfect and as such, the modeling process tries to incorporate the relevant cohort level heterogeneity and systematically build on the historical information, combined with business knowledge and prevailing industry trends. to provide fairly reasonable estimates of the future losses, integrating the experiences from both the crisis period that led to peak losses and the recent timeframes which has witnessed low losses due to the strong macro-economic environment.

Finally, the model has been tested on a multi-period forecast to ensure overall stability of the framework. This is primarily done given the path-dependent behavior associated with mortgages wherein single –period analyses do not always provide the best results leading to model over-fitting.

1.3 Vendor Models

[This section is applicable to vendor models only. Model Sponsor must provide the rationale for use of a vendor model instead of an internally developed model. The rationale could include at least the following factors

- the quality and transparency of the vendor model documentation
- the stability and integrity of the implementation

- the financial stability of the vendor
- the depth and transparency of the vendor model testing
- vendor model version controls and processes]

The 2019 redeveloped Method A RM model suite is not a vendor model.

1.4 Model Materiality Tier (MMT)

This section will provide additional information to the "Automated Document" Section A.2.: "Model Risk Rating" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

[Provide information required for MMT determination. Please refer to MRM policy for details on determination of MMT.]

The Model Materiality Tier (MMT) for the 2019 redeveloped Method A RM model suite is high, and is also included in the automated section at the beginning of this MDD.

1.5 Business Practices and Complexities

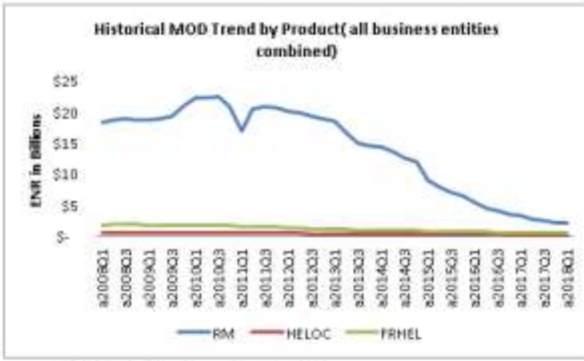
[Describe business practices and complexities that relate to the model's development and how the model mitigates, overcomes, and/or incorporates complexities?]

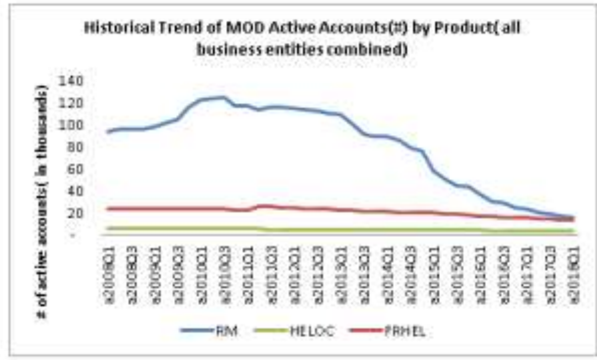
Presented below is the list of business complexities encountered during the process of RM model development –

Loan Modification (including trial modification) and Servicing released sale of mortgages

The loan modifications volume illustrated in figure 1.5.1 below pertain to loans modified for the North America real estate portfolio. Pure modifications are defined as the permanent restructuring of the loan. Under the permanent restructuring, one or more of the terms of the borrower's loan are changed to provide a more affordable payment option. This change can pertain to interest rate modification, term modification, principal balance reduction or a combination of any of these. The 2019 Method A group of models, segregates the modified loans from the non-modified loans and built separate models to account for the performance differences between these two subsets of loans. The MOD model considers the modified loans under its purview.

Figure 1.5.1: Modifications Trends (2008Q1 – 2018Q1)

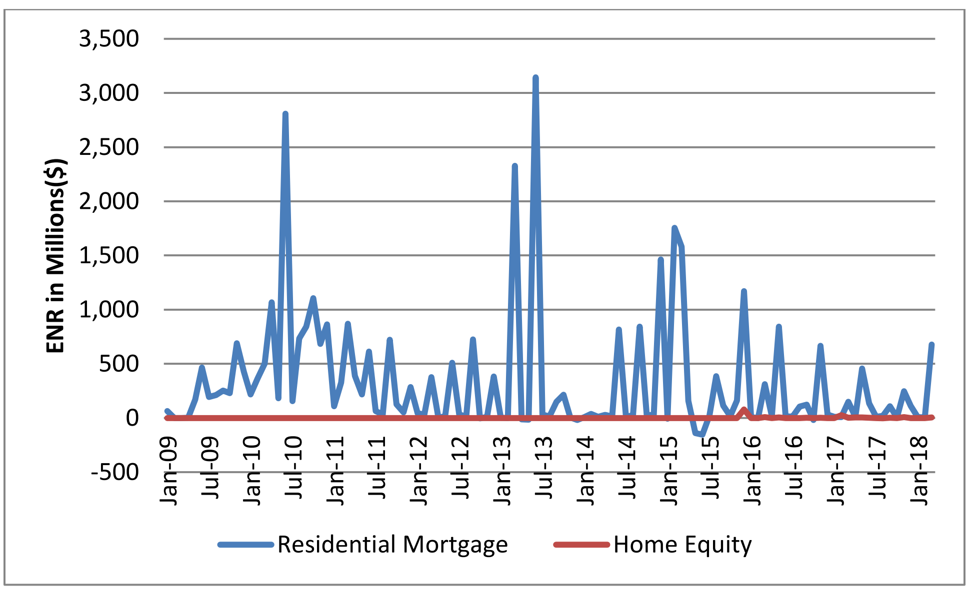




*2011Q1 Dip Reflects Penny Mac Asset Sales

An asset sale is defined as the non-recourse cash sale of assets from a bank or government agency to a third party. Asset sales are primarily management – led strategic initiatives aimed at increasing cash flow, reducing bad debt risk and/or liquidation of specific assets from the balance-sheet. Figure 1.5.2 illustrates the historical asset sale trends for the Bank Mortgage (CMI) portfolio. As shown below, asset sales peaked in the fourth quarter of 2012 with approximately $3.0Billion of assets sold.

Figure 1.5.2: Asset Sale Trends (2009Q1 – 2018Q1)



As mentioned above, both asset sales and loan modifications' affect model development and extensive research has been conducted to analyze the effect of these 'action' loans on the model's performance.

Typically, any action loan results in right censoring of performance data although it has no effect on model forecasting results or the model's sensitivity to macro-economic factors as the model forecast always assumes that there is no future 'action'. This aligns with both the modeling objective and the future business plan (Gating Principles Section 3.2) which has shown significant diminishing trends in both asset sales and loan modifications' in recent times.

In prior years, the model developers chose the random conditional censoring as the primary approach for 2016 and 2017 Method A (Champion) PD models. Under this approach, all 'action' loans were removed from the backtest population resulting in significant limitations in terms of inconsistent population sizes between model development sample and backtest. This produced incomprehensive model performance due to removal of action loan pre-action performance and discrepancy in backtest volume across different snapshots and compared to general ledger.
For the 2018 CCAR process, the model development team utilized two different approaches to handle the action loans: Naïve censoring & 2018 censoring. While the naïve approach removed both actual and predicted post-action performance for action loans, the 2018 censoring strategically redistributed the action loans' predicted liquidation probabilities prior to the action date, back to the active loan statuses. A quick comparative analysis between these two approaches demonstrated that while the Naïve censoring approach tended to inflate the pre-action prediction for the terminal statuses (i.e., IVP or VP) by not removing the action loan's pre-action termination probabilities. In contrast the 2018 Censoring avoided the inflation of terminal event prediction by setting the action loan's pre-action terminal probability at zero. However, MRM rejected the new approach as it was not consistent with the model development and forecasting as well as it was not a blind back-test. Hence, MRM rejected the new approach and approved the naïve approach as the final approach for back-testing. For the 2019 CCAR process, CAMU continued to adhere to MRM's directive and use the naïve censoring approach for all backtesting purposes.

Changing Mortgage Landscape and the idiosyncratic nature of the US real estate portfolio

There is likely no sector as important to the U.S. economy as housing. As of quarter one of 2018, Real estate and rental and leasing accounted for 13.4% of the total GDP. Including the spending on housing services as well as spending on various kinds of housing construction, the entire real estate industry can account for as much as one fifth of overall output in the U.S. economy. And this is why housing has traditionally powered the American economy out of recessions, and that's why housing's role as the trigger of the Great Recession was of such mammoth proportions.

The US economy had witnessed a steady increase in housing prices since 2000's. The housing prices peaked sometime around 2005-2007 with 50% of the real estate transactions allocated to subprime originations. Just like any other major banking organization, Bank also held a considerable part of its portfolio in subprime loans at this time with a fair proportion of them being originated through the third party (Wall Street), broker and correspondent channels through secondary marketing. Once the Financial crisis hit, Bank had to incur substantial losses, given the riskier concentration of its portfolio mix. The net credit losses at the peak of the crisis hovered around \$550MM- \$700MM per quarter exclusively for the BankMortgage portfolio. In the aftermath of crisis, a series of regulations and laws were passed, of which the most notable ones were the CCAR Stress Testing and Dodd-Frank Acts. Not only were new regulations implemented, but new regulators like the Consumer Financial Protection Bureau were created.

At the same time, the CFPB and other agencies became more assertive in their enforcement practices. Given the heightened expectations from the regulators, Bank created a strategic plan to systematically reduce its portfolio risk and initiate a change in its portfolio composition mix either through –

- Completely exiting out of third party/broker led origination channels, such as Wall Street
- Creating Bank Holdings which acted as a SPE(Special Purpose Entity) of sorts to house the non-performing loans that were relegated to periodic asset sales, given the evolving market conditions
- Focusing on retail originations (loans originated through Bank's branches) with stringent underwriting criteria and stricter home loan eligibility requirements.

The afore-mentioned strategic foot-prints led a considerable downsizing of Bank's Risk portfolio (trends and numbers in the next section) and a significant shift in the portfolio composition. A more detailed description of the individual portfolios is presented in the next section.

Post the financial crisis, given the effective oversight of external regulators, improving macro-economic conditions and Bank's self-initiation of a zero-tolerance/ 100-percent-compliance regime towards risk-management, helped lower the portfolio losses dramatically over the recent past.

As discussed above new regulations as well as Bank's own strategic shift in originating and managing mortgage portfolio resulted in Losses went down from $673MM at the peak of the crisis to ~$6-7MM in recent times.

Loss Trend in Recent Times

Another prominent feature associated with the afore-mentioned portfolios are the notable shifts in their loss pattern in the recent past accentuated with low volume and pronounced volatility. The charts presented below evidence of the considerable shifts in loss pattern since the historical stress and the observed volatility in recent period losses.

Figure 1.5.3: Loss (NCL) Trends Comparison - Recent Vs Historical Stress [CMI - RM Only]

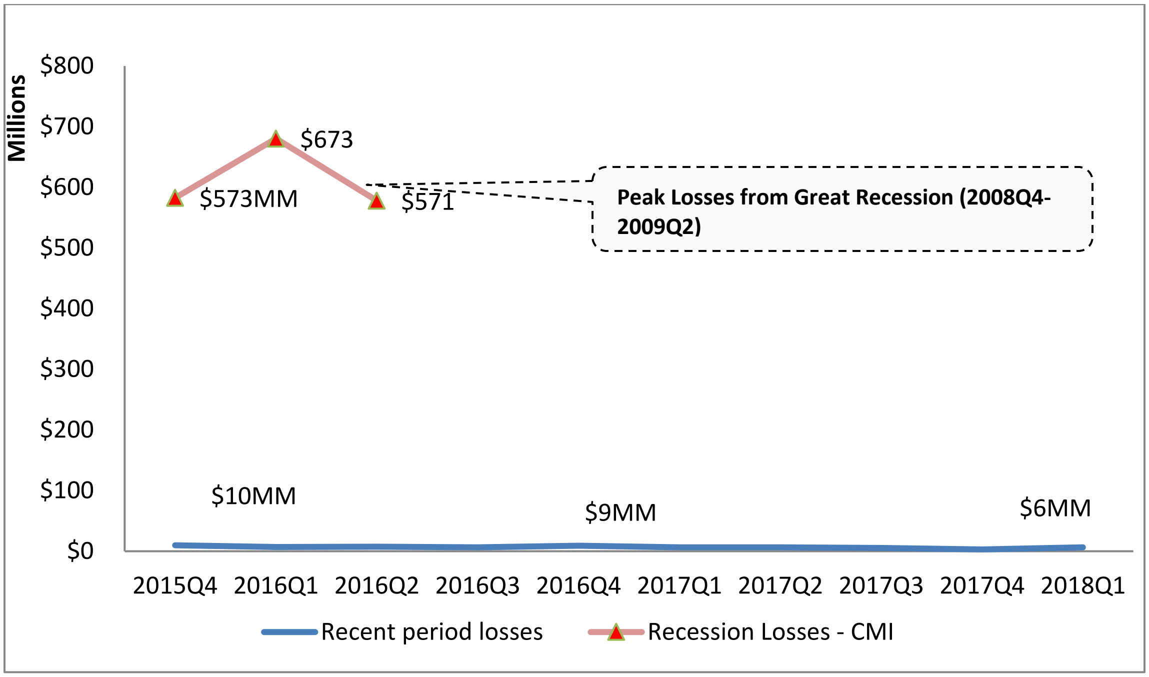
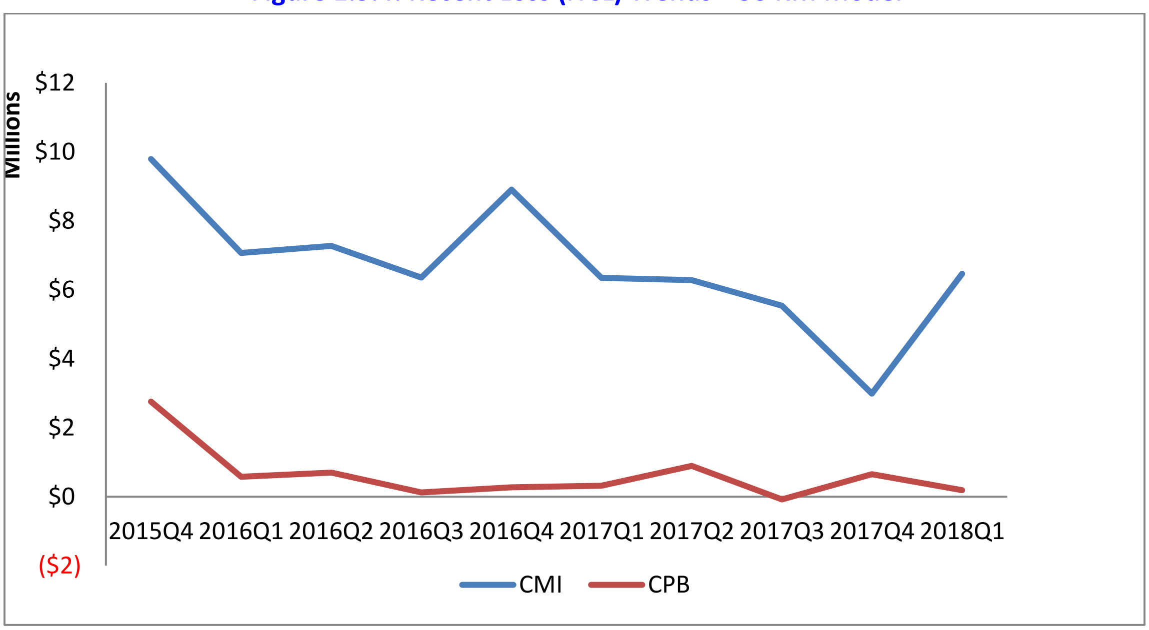


Figure 1.5.4: Recent Loss (NCL) Trends - US RM Model



The lower levels observed for recent period losses are a combined result of three critical drivers described below:

1. The credit profiles of today's portfolios are significantly better than what existed prior to the recession in 2007 given tighter underwriting standards and proactive portfolio management actions taken in recent years

2. Loans that were not performing as anticipated have received additional focus and/or treatments and as a result are no longer in the portfolio(s). Examples include short sales, strategic asset sales, revised write-off policy, and coupled with surviving borrowers that continue to prove their resilience as the bank has enhanced communications and its commitment to responsibly providing financial services that enable growth and economic progress.
3. Stronger macro-economic conditions in recent years fostered through increasing rates, rising GDP and housing prices and lower unemployment which additionally fuels the purchasing power of the borrower leading to much lower and infrequent losses

Given the fact that bank finances and risk management policies are much stronger and data more reliable in recent years, it creates somewhat of a paradox when assessing the quantitative accuracy of stress tests for these recent periods and their reasonableness via-a-vis the Great Recession experience. The new model is able to capture the secular trend in both the stress period as well as recent phase of strong portfolio performance due to low loss rate trend. Additional discussion on this can be found under 'Model Performance'.

Government Loans

There has been an influx in loans transitioning from BUK7 to IVP in recent times (2016+), especially for the government loans. This trend however contradicts with the broader macroeconomic outlooks which are marked with significant and consistent home price appreciation and unprecedented low unemployment rates. This peculiar data anomaly is partly influenced by the foreclosure process specific to government loans (if Bank wins the bid during the foreclosure sale, the loan tends to stay active within Bank's systems as an IVP loan even though the loan technically is no longer live). The second reason for the uptick in the IVP rates can be attributed to the personnel shifts which led to instability in management process and subsequent issuance of MRA's from OCC around the management of debts previously contracted (DPC). Since both of these factors arise from both operational inefficiencies and policy regime, it creates a modeling challenge to systematically model the IVP rates for the government loans. Additional details around this have been discussed in Section 4.1.5 of the MDD.

IVP Prediction Challenge

An involuntary payoff happens as the last resort for borrowers who are unable to meet their monthly mortgage obligation and the lender initiates the proceedings to take possession of the property. Once the loan hits the IVP status, the next steps of action and the loan's subsequent status relies less on the borrower's credit quality or loan performance and more on the given company's foreclosure and loss mitigation policies and the state's (loan's state of origination) regulation around the foreclosure process. Loans which fall under the purview of state-mandated foreclosure laws or have been granted relief as per the Foreclosure Moratorium Acts (**https://www.Bankgroup.com/Bank/news/2009/090213a.htm**,**https://blogs.wsj.com/washwire/2009/02/13/chase-Bank-to-implement-foreclosure-moratoriums/?mod=article_inline)** tend to experience a prolonged stay in the IVP status, and are insensitive to the borrower's changing credit profile or the broader macro-economic factors. Additional details around this will be discussed in Section 5.4 of the MDD.

Impact from Interest Rate Volatility

Please refer to MDD sections 2 and 6 below for extensive discussions on this topic.

1.6 Review and Challenge Process

[Provide a summary of the review and challenge process and discussions between Model Developers, Model Sponsors and Independent Risk that happened throughout the model development process. Also, provide evidence of detailed discussions.]

There are a significant number of stakeholders that participate in CCAR model development, effective challenge, and validation activities throughout the model lifecycle.

Key in-business stakeholders involved in effective challenge of risk models, which includes 1st line of defense assessment of conceptual soundness include:

- **In-Business Chief Risk Officer (CRO):** Oversees all aspects of Risk Management for respective line of business
- **Model Sponsor:** Responsible for all decisions related to the model during the model lifecycle including adherence to policy requirements at every stage and attests to model inventory semi-annually
- **Regional Modeling Head (RMH):** Identifies modeling needs and oversees / directs modeling related activities in respective region or business group. The RMH ensures oversight and governance of the model throughout the model lifecycle and is responsible for all model administrative matters in relation to global policy
- **Model Developer**: Builds models consistent with the functional/technical requirements defined by the model sponsor and model stakeholders, ensuring that the models comply with all global policy and have in-business maker/checker functions throughout the Model lifecycle
- **Credit Policy, Model End Users, and Senior Risk Management:** Serve as in-business Business Peers / Sponsors that review the model framework and results, offer guidance, and *effectively challenge* the model development team throughout the model lifecycle.

Key independent stakeholders involved which include 2nd line of defense assessment of conceptual model soundness include:

- **Independent Risk Modeling Oversight (IRMO)** provides guidance on global policy requirements and independently validates models for functional soundness
- **Model Risk Management (MRM)** provides guidance on global policy requirements and independent validates models for technical soundness

- **Fair Lending, Compliance, and Legal Counsel** provides guidance and reviews the model documentation post independent validation and prior to placing model in use to ensure there are no regulatory constraints for the proposed model

Key evidence of stakeholder contribution is that model stakeholders participate in the following activities to provide input, review, and challenge:

- Annual Modeling Plan preparations
- Preparatory and checkpoint sessions to gather model end user feedback
- Weekly update provided to USCM Chief Risk Officer
- IRMO/MRM/IROC meetings

All supporting documentation and contributions specific to the proposed new model, including detailed stakeholder meeting minutes, are included as evidence of effective challenge below.

- CAMU Annual Modeling Plan – 2018
  CCAR Remediation Plan + Model Sponsor Approval + MRM Approval (See attachments' subfolder – '1.6 CAMU Annual Modeling Plan – 2018').

- Model End User Review Meeting Minutes (See attachments' subfolder – '1.6 Model End User Review Meeting Minutes').

- Model Risk Management (MRM) Meeting Minutes (See attachments' subfolder – '1.6 Model Risk Management (MRM) Meeting Minutes').

- Independent Risk (IROC/IRMO) Effective Challenge Presentation and related meeting minutes, Review Memorandum, & Follow-Up Tracking (See attachments' subfolder – '1.6 Independent Risk (IROC_IRMO) Effective Challenge ').

MRM Question - The model name has been updated from 1st Lien model to Residential Model as compared to the previous version of the model. Sponsor need to confirm if the scope of this model is the same as the previous model. Also confirm if the portfolio covers the 1st lien mortgage, and 1st lien HELOC and FRHEL. Any differences should be provided. 2nd MEA: This is not provided. Model Sponsor is requested to provide the same.

Answer – The scope of the 2019 Method A RM model extends to both CMI and CPB residential mortgage loans across all liens. These include both first and 2^{nd}+ liens. Home equity loans and lines of credit are outside the purview of the RM model Specifically the home equity Lines of credit (HELOC) are considered with the 2019 Method A HELOC model while the fixed rate home equity loans are considered within the 2019 Method A FRHEL model. Please see Section 1.2 for additional details.

MRM Question - Portfolio Mix Change

It is observed that CMI portfolio is reducing over time and CPB portfolio is increasing. Model Sponsor needs to provide details on the portfolio mix change and how this model is addressing this portfolio mix change. This was also asked by regulators during the prior examinations.
This information is not provided. Model Sponsor is requested to provide detail analysis and assessment about the portfolio mix change.

Answer – Please note that the Gating Principles' cited within Section 3.2 describes the idiosyncratic nature of each modeled portfolio(CMI, CPB). Please note, while there has been a distinct shift in portfolio composition over time [CMI portfolio reducing over time and CPB increasing over time], CAMU believed such a shift should not impact the model performance because –

a) The model is built at a loan level and not at a portfolio level. As such all idiosyncratic loan level attributes have been considered as part of the model development process

b) While the CPB portfolio is increasing in terms of overall volume, please note that the absolute number of CPB loans is not significant compared to CMI.
For other pertinent details please refer to Section 3.2 of the MDD.

2. Limitations and Compensating Controls

This section will provide additional information to the "Automated Document" Section A.4.: "Limitations" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

[Clearly list the self-identified limitations of the model and identify cases where the model should or should not be used. These cases must consider the overall objective of the model and be tangible and enforceable, preferably at the system level. Examples include model/method combinations for particular products and underlying type restrictions (e.g. roll rate models cannot be used for stress testing purposes of credit card portfolios). Also, mention the limitations and weaknesses identified by the validator in the prior version of the model, if applicable. Describe how each of these limitations were mitigated/ can be mitigated such as through compensating controls on the model.]

This chapter identifies cases where the Model should or should not be used. These cases consider the overall objective of the model, the tangibility/enforceability, of the proposed modeling approach and application of all inputs and outputs at the appropriate segment level. The proposed model, if approved, would be tracked and maintained as an End User Computing (EUC) tool. Chapter 7-9 of this MDD includes all management oversight and access controls for the proposed model, along with User Guide, User Acceptance Testing, and Implementation Testing evidence. The residential mortgage loans in scope include loans held in portfolio as well as loans Serviced for Others (SFO). The loans encompass both the CMI and CPB businesses respectively. Model usage, or any limitations thereof, are tracked and reported to Independent Risk and Model Risk Management as part of the Quarterly Model Sponsor Attestation process.

The critical model limitations corresponding to the 2017 Method A and 2018 Method B model suites and their corresponding remediation action plans have been discussed in detail in Section

1.1 of the MDD. Please refer to attachment – '2 Prior Limitations and Remediation Plan_2019.xlsx.' for a holistic overview of the open limitations from prior year's models (both Method A and Method B model suites) and how they were addressed in this year's modeling efforts. The below table includes all prior model limitations noted that impact the RM portfolio as noted in Champion and Challenger model validation reports (i.e. 1st Lien, DVM, Severity), along with corresponding limitation ID, required action, compensating control, and proposed remediation for 2019 CCAR submission.

Table 2.1: Prior 1st Lien Model Limitations Remediation Summary

Method A – 1st Lien (Residential Mortgage) Technical Limitations

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
16575	**Backtesting** CCAR usage – For some final forecasted variables and/or individual transition equations, the model breached trigger in the non-stress and/or stress period backtest for the medium term forecast horizon. There were instances where the model was not breaching the trigger but was under predicting the results.	For the CCAR usage, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCARGCRM/2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
16576	**Backtesting Approach** The new evaluation backtest approach adjusts the model output and is not a blind backtest.	MRM restricts the usage of the new evaluation backtest approach. Model Sponsor should use the blind (naïve) approach as the final backtest approach.	All future Quarterly Performance Assessments (QPAs) will comply with the CCAR Loss Forecasting Model Performance Testing Guidance. Specifically, the quarterly back test results will be reported to and evaluated by the Model Sponsor, Loss Forecasting, Portfolio Management, Model Risk Management, and Independent Risk – utilizing the blind (naïve) approach. All Quarterly Performance Assessments (QPAs)	All backtesting results were calculated utilizing the blind/naïve approach which were included in the 2019 CCAR MDD submissions commensurate with MRM's Loss Forecasting Model

			are housed in iMRMS.	Performance Testing Guidance.
21795	**Backtesting** Non CCAR usage - For some final forecasted variables and/or individual transition equations, the model breached trigger in the non-stress and/or stress period backtest for the short, medium and/or long term forecast horizon. There were instances where the model was not breaching the trigger but was under predicting the results.	For the non CCAR usages, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCARGCRM/2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
16577	**Model Performance** 1. For some final forecasted variables and/or individual transition equations, the model breached trigger in the non-stress and/or stress period backtest for the short, medium and/or long term forecast horizon. There were instances where the model was not breaching the trigger but was under predicting the results. 2. Redevelopment of the model was not considered given the major changes in the model input data: a) New HPI variable (v4) b) Changes in historical CPB loss data c) Update of one-time loan list Inclusion of all loans in the	MRM recommends redevelopment/replacement of the model. Hence, Model Sponsor is required to submit the redevelopment/ replacement plan to MRM. This plan should also consider observations made by Regulators and the validation report.	CAMU will re-develop the Method A 1st Lien/RM model and attempt to address all segment/transition level triggers, especially as it pertains to any under prediction; Method A MDD submissions are targeted to be delivered to MRM by 9/30/18. Method A is a probabilistic framework that uses weighted average variables to calculate discrete probability delinquency counter variables. Method B addressed item (4) through simulation framework, and will further discuss and document interpretations of delinquency counter variable in Method A.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Specifically, root-cause analysis was performed on trigger breaches, changes in input data were evaluated, along with variables subject to interpretation was all considered in the redevelopment process. The proposed enhancements, data, modeling approach, final variables, and outcomes analysis for the new model are highlighted in the Executive Summary and detailed in Chapters 3, 4, 5, and 6

	modified code logic without filtering out the warehouse held for sale. 3. Under prediction of losses for higher CLTV bands was observed 4. Model is using variables, such as delinquency counter variable, which are subject to interpretation during implementation and can't be implemented the same way as used in development.			of the MDD.
16578	**Alternative Modelling Approach** MRM notes that both Method A and B approaches use similar modelling framework (transition matrix framework) that could not be viewed as real alternatives.	MRM recommends redevelopment/replacement plan to include these points. Sponsor must explore alternate modelling techniques.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A 1st Lien/RM model. As an additional primary focus, CAMU is developing a Method C model suite with a hazard modeling approach which will also include a 1st Lien/RM model. Proposed Method C timeline includes sharing 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Although the Method A 1st Lien/RM utilizes a transition approach, the Method C suite leverages a completely different modeling framework to mitigate the prior models' limitations around similarity of the modeling framework. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.
16579	**Modelling Approach Complexity** Transition equation model framework is a complex framework for stated objective and usage. Complexity creates problem during development/redevelopment, diagnosis, implementation,	MRM recommends redevelopment/replacement plan to include these points.	It is important to note that the transition framework has the following purpose(s): - Model is built to serve multiple objectives - BAU, CECL, Business Planning, TDR that require monthly loan level lifetime loss forecast. - Balance between number of models for multiple objectives vs complexity of model	The Method A transition model approach is in the process of being benchmarked to the alternative hazard approach Method C suite. The Method C suite leverages a

	model testing, documentation etc. MRM observes that the modeling framework can be simplified and the objective can be achieved using simpler approaches.		suite. - Transition model has been a commonly used credit risk modeling framework and exists to forecast both intermediate and terminal status of a loan - MRM has commented on addressing key risk segment difference in the new model development. To address this limitation, additional complexity will have to be introduced such as interaction variables and split certain equations. In order to fully address this model limitation noted by MRM, CAMU is also developing a Method C model suite with a hazard approach which will attempt a less complex methodology for 1st Lien/RM product portfolios. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19.	completely different modeling framework which is less complex than the transition approach. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.
16580	**<u>Model Benchmarking</u>** The CCAR 2018 Method A model was benchmarked with the previous version of the same model (CCAR 2017 Method A) which has the same model co-efficient and structure. Alternate modelling approaches were not used for the purpose of benchmarking analysis.	Model Sponsor is required to provide the benchmarking analysis with alternate modeling approach.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A 1st Lien/RM model. As an additional primary focus, CAMU is developing a Method C model suite with a hazard modeling approach which will also include a 1st Lien/RM model. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Additionally, Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how to best present all uses, benchmarking of methods A/B, and granular level characteristic analysis. This will be updated beginning with 1Q18 MIS (published in 2Q18).	The Method A transition model approach is in the process of being benchmarked to the alternative hazard approach Method C suite. The Method C suite leverages a completely different modeling framework to mitigate the prior models' limitations around similarity of the modeling framework. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.
16581	**<u>Documentation</u>** MRM observed the documentation was insufficient at some places including mentioned below	Model Sponsor is required to provide complete documentation.	Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how to best present all uses, benchmarking of methods A/B, and granular level	The Method A transition model approach is in the process of being benchmarked to the

	1. The appropriateness of combining CPB with CMI and usage of the CMI model for SFO portfolio was not provided. 2. Complete benchmarking analysis as per the model usage was not conducted. 3. Granular level characteristics analysis was not provided. Also, characteristics analysis for the most recent period was not conducted.		characteristic analysis. This will be updated beginning with 1Q18 MIS (published in 2Q18), and will also be updated in annual Model Review (AMR) submissions for existing Method A model suite, along with Model Development Documents (MDDs) for new Method A suite on target to be submitted by 9/30/18.	alternative hazard approach Method C suite. The Method C model suite (due for submission 1Q19) uses a hazard modeling framework, which is substantially different from the transition model framework. The hazard framework models the termination events in a competing risk framework. In other words, for each month of a scenario forecast, the model's competing risk survival methodology predicts the probabilities of the loan termination events – ICP, VP first time BUK7 and loans that are either sold & released or sold & serviced, conditional on the survival of the account to date. As such, benchmarking analyses, backtesting analyses, characteristic analyses, and support for CMI/CPB coping strategy will be included in the MDD submissions.
16582	**<u>Code and Data</u>** Multiple challenges were faced during the independent testing of the model including incomplete/missing codes and datasets, access issues and mismatch in results.	Model Sponsor is required to adhere to the code and data guidance so that these issues can be minimized.	CAMU worked daily with MRM during validation to clarify code/datasets. Annual Model Review (AMR) submissions for existing Method A model suite, along with Model Development Documents (MDDs) for new Method A suite will be checked for completeness, accuracy, and replicability in compliance with MRM Data Guidance, and are being prepared for submission by 9/30/18. CAMU employs a maker checker effective challenge process for all new model development as well as ongoing model monitoring activities.	Code implementation and code execution for report generation were carried out by separate sub-teams within CAMU's production team, along with a third step of independent verification of final reports prepared from model outputs. This approach allowed multiple rounds of verification of the code package and enabled the team to identify updates needed to

				prepare the code package for delivery to independent validation (MRM) and business user teams. CAMU uploaded the proposed model data and code package to the CCAR_MVG directory in compliance with MRM's Code and Data Guidance.

<u>Method A 1st Lien (Residential Mortgage) – Functional Limitations</u>

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
19638	**<u>Model Performance</u>** The model does not seem to be sufficiently sensitive. The model under predicts in the stress period and over predicts in the base period. In addition, the stress multiples on PD are lower than expected for higher quality loans, some non-modeled assumptions may be countercyclical, and the partial charge off logic may tend to over predict. Overlays should be considered given the uncertainty around explaining the recent period back test performance. In addition, the analysis should consider the sales and modification activity.	The Sponsor should consider overlays to account for the back testing breaches and under-prediction. The sponsor should also consider the impact of model assumptions in given macro scenarios, and should consider the weaknesses in the PCO logic. In addition, the model results should be benchmarked against other models.	Instances where under/over prediction was observed in end to end loss testing was addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCARGCRM/2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in end to end NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.

19775	**Modeling Approach** Future development should consider: 1. Additional segmentation for CMI and CPB, product types, and risk drivers. 2. Additional review of potential risk drivers and macroeconomic relationships. 3. Research on making the non-modeled transitions more dynamic or at least prevent them from being countercyclical. 4. Enhancing the partial charge off logic.	The redevelopment plan that the sponsor is required to submit to MRM, should also address these observations.	Method A 1st Lien/RM model redevelopment will attempt to address items 1-4. Model redevelopment will include detailed segmentation analysis for various business portfolios, products, and risk drivers. Partial Charge off logic is being enhanced. Method A MDD submissions are targeted to be delivered to MRM by 9/30/18. Additionally, CAMU has confirmed with Finance/Accounting the exact accounting rules (partial charge off policy) for loans staying in bucket 7, curing from bucket 7; to lower delinquency bucket and second time entering bucket 7. Finance/Accounting confirmed: 1) Collateral dependent loan will be assessed a write-down once every year, 2) other FFIEC loans will be assessed every month as long as they are >=60dpd. (Other FFIEC loans means loans in FFIEC file but not collateral dependent. A cured loan will start to be subject to write-down when it is >=60dpd); and 3) If PCO>UPB, then FFIEC will receive a credit, i.e. PCO will decrease to UPB.	The model development team developed a systematic approach based on historical data movements, key risk driver movements and macro-economic trends to empirically rationalize the intuition behind modeled and non-modeled transitions to account for the counter-cyclical movements around certain transitions. To add context, as part of the 2019 PD model redevelopments all transitions were evaluated to distinguish and separate out the modeled vs modeled transitions using a statistical waterfall approach that combined the transition volume, contribution volume between transitions, volatility associated with each transition along with their measured C-statistics to deem their eligibility for being modeled as equations. For a transition that was deemed as non-modeled, an empirical lookup logic was created that accounted for the base vs stress, in-trial vs not-in trial and government vs conventional differentials to account for their differentiated performances across the afore-mentioned parameters. Further, this lookup logic was tested using a 10% shock to illustrate the model sensitivity to these assumptions. As such, segmentation analyses, CMI/CPB coping strategy, enhanced PCO logic, and support for non-modeled transitions were thoroughly documented in the MDD

				submission.

<u>Method B 1st Lien (Residential Mortgage) – Technical Limitations</u>

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
16735	**<u>Backtesting</u>** CCAR usage – For some final forecasted variables and/or individual transition equations, the model breached trigger in the non-stress and/or stress period backtest for the medium term forecast horizon. There were instances where the model was not breaching the trigger but was under predicting the results.	For the CCAR usage, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCARGCRM/2018CCAR/SitePages/Home.aspx It is important to note that the Method B model suite was not utilized in the CCAR forecast; only as a benchmark.	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
21799	**<u>Backtesting</u>** Non CCAR usage - For some final forecasted variables and/or individual transition equations, the model breached trigger in the non-stress and/or stress period backtest for the short, medium and/or long term forecast horizon. There were instances where the model was not breaching the	For the non CCAR usages, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCARGCRM/	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period

	trigger but was under predicting the results.		2018CCAR/SitePages/Home.aspx It is important to note that the Method B model suite was not utilized in the CCAR forecast; only as a benchmark.	snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
16736	**Model Performance:** 1. For some final forecasted variables and/or transitions, the model breached trigger in the non-stress and/or stress period backtest for the short, medium and/or long term forecast horizon. There were instances where the model was not breaching the trigger but was under predicting. 2. No consideration of recent 2 years of data either for model development or model stability analysis. 3. Only a few equations were redeveloped. Model was partially redeveloped. 4. Backtest approach is not completely blind. It uses future 'known' action rate information. 5. Model introduces additional complexity by converting intermediate probabilities into discrete events using simulation approach without any significant advantages. 6. Some independent variables' co-efficient signs appear counterintuitive. 7. Grouping of transitions was not thoroughly tested/justified in terms of key risk drivers of grouped	MRM recommends redevelopment/replacement of the model. Hence, Model Sponsor is required to submit the redevelopment/ replacement plan to MRM. This plan should also consider observations made by Regulators and the validation report.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which will include which will include sufficient documentation of items 1-8 in Method A 1st Lien/RM model; Method A MDDS are targeted for delivery to MRM by 9/30/18. Items 1-8 will also be addressed in Method C new model development, and CAMU will share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission, and items 1-8 were considered in during redevelopment for the new proposed model. Specifically, root-cause analysis was conducted to evaluate all back testing errors. More recent data was included in the development. Non modeled and modeled transitions were thoroughly evaluated during the development. Back test was completed utilizing blind/naïve approach. The proposed enhancements, data, modeling approach, final variables, and outcomes analysis for the new model are highlighted in the Executive Summary and detailed in Chapters 3, 4, 5, and

	buckets. 8. Model under predicts losses for higher CLTV bands.			6 of the MDD.
16737	**Alternative Modelling Approach** 1. MRM notes that both Method A and B approaches use similar modelling framework (transition matrix framework) that could not be viewed as real alternatives. 2. Alternative segmentation scheme was not explored and tested.	MRM recommends redevelopment/replacement plan to include these points. Sponsor must explore alternate modelling and segmentation techniques.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A 1st Lien/RM model. As an additional primary focus, CAMU is developing a Method C model suite with a hazard modelling approach which will also include a 1st Lien/RM model. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A transition model approach is in the process of being benchmarked to the alternative hazard approach Method C suite. The Method C suite leverages a completely different modeling framework. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.
16755	**Modelling Approach Complexity** Transition equation model framework is a complex framework for stated objective and usage. Complexity creates problem during development/redevelopment, diagnosis, implementation, model testing, documentation etc. MRM observes that the modeling framework can be simplified and the objective can be achieved using simpler approaches.	MRM recommends redevelopment/replacement plan to include these points.	It is important to note that the transition framework has the following purpose(s): - Model is built to serve multiple objectives - BAU, CECL, Business Planning, TDR that require monthly loan level lifetime loss forecast. - Balance between number of models for multiple objectives vs complexity of model suite. - Transition model has been a commonly used credit risk modeling framework and exists to forecast both intermediate and terminal status of a loan - MRM has commented on addressing key risk segment difference in the new model development. To address this limitation, additional complexity will have to be introduced such as interaction variables and split certain equations. In order to fully address this model limitation noted by MRM, CAMU is also developing a Method C model suite with a	The Method A transition model approach is in the process of being benchmarked to the alternative hazard approach Method C suite. The Method C suite leverages a completely different modeling framework which is less complex than the transition approach. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.

			hazard approach which will attempt a less complex methodology for 1st Lien/RM product portfolios. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19.	
16775	**Model Benchmarking** The CCAR 2018 Method B model was benchmarked with the previous version (CCAR 2017 Method B) of the model which uses the same modelling approach and was not completely re-developed (only few equations were developed). Alternate modelling approaches were not used for the purpose of benchmarking analysis.	Model Sponsor is required to provide benchmark analysis with alternative modelling approach.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A 1st Lien/RM model. As an additional primary focus, CAMU is developing a Method C model suite with a hazard modelling approach which will also include a 1st Lien/RM model. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A transition model approach is in the process of being benchmarked to the alternative hazard approach Method C suite. The Method C suite leverages a completely different modeling framework. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses will be included in the Method C MDD submission.
16776	**Documentation** MRM observed the documentation was insufficient at some places including mentioned below: 1. Appropriateness of combining CPB with CMI and usage of the CMI model for SFO portfolio unavailable. 2. Usage of the CMI stress factor for the CPB LGD model was not justified. 3. Justification of the loan level accuracy of simulation approach. 4. Complete benchmarking analysis per model usage was	Model Sponsor is required to provide complete documentation.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which will include which will include sufficient documentation of items 1-9 in Method A 1st Lien/RM model; Method A MDDS are targeted for delivery to MRM by 9/30/18. Items 1-9 will also be addressed in Method C new model development, and CAMU will share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission, and items 1-9 were considered in during redevelopment for the new proposed model. Specifically, documentation was enhanced with sufficient rationale for CMI/CPB/USRB coping strategy, justification for loan level accuracy, segmentation analysis and variable justification, as well as data treatments. The proposed

	not conducted. 5. Granular level characteristics analysis unavailable. 6. Incomplete justification and analysis about including historical delinquency variables. 7. Correlation analysis was conducted using the HPI v3 version and not the HPI v4 version. 8. For the key risk drivers such as FRM vs ARM, the sensitivity analysis was conducted only for the initial transitions buckets and the analysis on the other buckets was not shown. 9. Complete details on missing data treatment/capping/flooring were unavailable.			enhancements, data, modeling approach, final variables, and outcomes analysis for the new model are highlighted in the Executive Summary and detailed in Chapters 3, 4, 5, and 6 of the MDD.
16777	**Code and Data** Multiple challenges were faced during the independent testing of the model including incomplete/missing codes and datasets, access issues and mismatch in results.	Model Sponsor is required to adhere to the code and data guidance so that these issues can be minimized.	CAMU worked daily with MRM during validation to clarify code/datasets. Annual Model Review (AMR) submissions for Method B model suite will be checked for completeness, accuracy, and replicability in compliance with MRM Data Guidance, and are being prepared for submission by 9/30/18. CAMU employs a maker checker effective challenge process for all new model development as well as ongoing model monitoring activities.	Code implementation and code execution for report generation were carried out by separate sub-teams within CAMU's production team, along with a third step of independent verification of final reports prepared from model outputs. This approach allowed multiple rounds of verification of the code package and enabled the team to identify updates needed to prepare the code package for delivery to independent validation (MRM)

				and business user teams. CAMU uploaded the proposed model data and code package to the CCAR_MVG directory in compliance with MRM's Code and Data Guidance

<u>Method B - 1st Lien (Residential Mortgage) – Functional Limitations</u>

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
19436	In spite of the improvement for this round of model development, the CPB loss data continues being a big challenge for the model development. This restricts the LGD methodology selection for the CPB portfolio and impacts the model performance for the CPB portfolio.	The Model Sponsor will need to continue working with the CPB business team on the CPB loss data in order to achieve better data completeness and data reliability. The LGD Modeling methodology will need to be improved when data allows.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which will include CPB loss data in Method A 1st Lien/RM model; Method A MDDS are targeted for delivery to MRM by 9/30/18. CPB loss data will also be closely scrutinized in Method C new model development, and CAMU will share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/15/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The CPB loss data was thoroughly reconciled as part of model redevelopment. The re-developed Method A transition approach, as well as the new proposed Method C hazard approach include more recent data, as detailed in MDD Chapter 4. It is important to note that the Method B model suite was not utilized in the CCAR forecast; only as a benchmark.
19437	This model uses simulation approach in the implementation instead of the probability multiplication approach. Although there is further analysis to support model result stability at more granular levels, it is not sufficient to support the reliability and stability of the loan level estimates.	The Sponsor needs to provide assessment and strategy for use of the loan level results if this model is used for any business applications which require the loan level results.	Simulation approach utilized was recommended by a previously MRM validator who is no longer with Bankbank. For the 2018 CCAR Model Method-B, the business sponsor (model end user) confirmed the model output is not used in official Non-CCAR submission (e.g., BAU/LLR forecasts), and does not alter/impact any of the Non-CCAR results. Method B model suite is used solely for benchmarking purposes. An alternative loan level hazard model is being developed (Method C Model suite) which will also include a 1st Lien/RM model. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being	Both the re-developed Method A transition approach, as well as the new proposed Method C hazard approach (alternative) included full re-development. It is important to note that the Method B model suite was not utilized in the CCAR forecast; but rather used only as a benchmark.

			prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature. benchmarking purposes based on its simulation feature.	
19535	The model development started from 2017 Method B model, with full development only for the three key PD equations and the LGD model. However, the model backtest by risk attributes, together with the observations for the LGD methodology and the PD methodology suggests potential benefit from a more complete re-development.	The Sponsor should consider a full model development and should also explore alternative modeling approaches.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A 1st Lien/RM model. CAMU is testing an alternative modeling approach to simplify the two-stage severity model structure, and also testing as part of Method C Model development. CAMU is developing an alternative to B - which includes a Method C model suite with a hazard modeling approach which will also include a 1st Lien/RM model. Proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Although the Method A 1st Lien/RM utilizes a transition approach, the Method C suite leverages a completely different modeling framework to mitigate the prior models' limitations around similarity of the modeling framework. The forecast between Methods A and C will decompose the difference between the respective model results and generate a meaningful comparison. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses a will be included in the Method C MDD submission.
19536	For the benchmark analysis, although there is some information about Method A in the benchmark analysis, there is no clear comparison of the setup of backtest or forecast between Method A and Method B. As a result, it is hard to decompose the difference between their model results and make a meaningful comparison.	The Sponsor should set up the backtest for these two methods in a comparable way. Also, given the similarity in model approach between Method A and Method B, the sponsor should also explore alternative approach or industry forecast in order to help assess the reasonableness of the model results.	Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how various model results can be compared. This will be updated beginning with 1Q18 MIS. Method A MDD submissions are targeted to be delivered to MRM by 9/30/18. An alternative benchmark is being developed, and the proposed Method C timeline is to share 1st Lien/RM benchmark results by 12/15/18 and submit MDD by 1/31/19. Finally, Method B annual model renewal documentation is being prepared for submission by 9/30/18, and will continue to be used for benchmarking purposes based on its simulation feature.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Although the Method A 1st Lien/RM utilizes a transition approach, the Method C suite leverages a completely different modeling framework to mitigate the prior models' limitations around similarity of the modeling framework. The forecast between Methods A and C will decompose the difference between the respective model results and generate a meaningful comparison. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses a will be included in the Method C MDD submission.

19435	The model development didn't use the most recent two years of data when it is available. Although it allows the developer to perform the out of time validation test, the model may not be able to capture the change in the loss behaviour in the recent years.	The Model Sponsor needs to assess the impact from this model decision on the model results.	Annual Model Review (AMR) documentation for Method B model suite is planned for submission by 9/30/18. CAMU employs a maker checker effective challenge process for all ongoing model monitoring activities (including population stability index reviews). These results are discussed in Model Governance Risk Committee (MGRC) and minutes are available for MRM's review on CAMU's SP site. Quarterly Performance Monitoring evidence is also uploaded into iMRMS, and loss behavior in recent years will be closely monitored.	The re-developed Method A transition approach, as well as the new proposed Method C hazard approach include more recent data, as detailed in MDD Chapter 4. It is important to note that the Method B model suite was not utilized in the CCAR forecast; only as a benchmark.

DVM – Technical Limitation

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
14921	**Documentation** 1. Complete details on the benchmark model were not provided. 2. The model output was not tested for sensitivity to underline assumptions.	Model Sponsor is required to provide the complete analysis and documentation.	CAMU plans to redevelop the Distressed Valuation Model (DVM). A haircut logic will be developed and built into the new Method A model suite and documented in the respective MDD targeted for delivery by 9/30/18. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative. Independent Risk is setting up a working group to address ongoing model monitoring, of which benchmarking of methods A/B, and granular level characteristic analysis for underlying assumptions will be discussed. This will be updated beginning with 1Q18 MIS (published in 2Q18).	A new distressed value (DV) haircut logic has been developed and proposed for the Method A model suite. The new proposed DV logic is explained in detail within the HELOC MDD, which includes sensitivity analysis as required by MRM's Loss Forecasting Model Performance Testing Guidance. Benchmarking analyses a will be included in the Method C MDD submission. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative.

DVM – Functional Limitation

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation

19517	**Model Monitoring** The current model testing and monitoring does not provide a complete view of how the model errors impact the projected end to end losses.	Update monitoring to include metrics that provide a more complete view of how model errors impact the projected losses, such as the error in predicting the distressed value haircut.	CAMU plans to redevelop the Distressed Valuation Model (DVM). A haircut logic will be developed and built into the new Method A model suite and documented in the respective MDD targeted for delivery by 9/30/18. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative.	A new distressed value (DV) haircut logic has been developed and proposed for the Method A model suite. The new proposed DV logic is explained in detail within the HELOC MDD, which includes end to end sensitivity analysis as required by MRM's Loss Forecasting Model Performance Testing Guidance. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative.
19516	**Modeling Approach** Future development should consider directly modeling the distressed value haircut, and should further explore specifications that include risk drivers and macroeconomic drivers.	The Model Sponsor is required to consider the point during future redevelopment.	CAMU plans to redevelop the Distressed Valuation Model (DVM). A haircut logic will be developed and built into the new Method A model suite and documented in the respective MDD targeted for delivery by 9/30/18. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative. Independent Risk is setting up a working group to address ongoing model monitoring, of which benchmarking of methods A/B, and granular level characteristic analysis for underlying assumptions will be discussed. This will be updated beginning with 1Q18 MIS (published in 2Q18).	A new distressed value (DV) haircut logic has been developed and proposed for the Method A model suite. The new proposed DV logic is explained in detail within the HELOC MDD, which explored additional risk drivers and macroeconomic drivers. If the new model is validated and approved as the Champion, then it is anticipated the existing model will be the alternative.

Severity– Technical Limitation

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation

14918	**Backtesting** The model under predicts the results and breaches the backtest threshold for the short term recent and stress period backtest for the VA segment.	For the CCAR usage, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools. consumer.Bankgroup.net/sites/CCARGCRM/ 2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
21815	**Backtesting** Non CCAR usage - The model under predicts the results and breaches the backtest threshold for the short term recent and stress period backtest for the VA segment.	For the non CCAR usages, the Sponsor is required to consider overlay that would account for the back-testing breaches and under prediction observed.	All back-testing breaches and instances of under prediction were addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools. consumer.Bankgroup.net/sites/CCARGCRM/ 2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
14919	**Model Benchmarking** The CCAR 2018 Method A model was benchmarked with the previous version of the same model (CCAR 2017 Method A) which has the same model co-efficient and structure. Alternate modelling approaches were not used for the purpose of benchmarking analysis.	Model Sponsor is required to provide the benchmarking analysis with alternate modeling approach.	The priority focus is re-developing the existing approved Champion Model Suite by 9/30/18, which includes Method A Severity model. As an additional primary focus, CAMU is developing a Method C model suite with a hazard modeling approach which will also include a Severity component. Proposed Method C timeline is to share Severity benchmark results by 12/15/18 and submit MDD by 1/31/19. Additionally, Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how to best present all uses, benchmarking of methods A/B, and granular level characteristic analysis. This will be updated beginning with 1Q18 MIS (published in 2Q18).	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Although the Method A 1st Lien/RM utilizes a transition approach, the Method C suite leverages a completely different modeling framework to mitigate the prior models' limitations around similarity of the modeling framework. The forecast between Methods A and C will decompose the difference between the respective model results and generate a meaningful comparison. Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (1Q19). As such, all benchmarking analyses a will be included in the Method C MDD submission.

14920	**Documentation** • The appropriateness of combining CPB with CMI/USRB was not provided. • The model output was not tested for sensitivity to underline assumptions.	Model Sponsor is required to provide the complete analysis and documentation.	Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how the model errors impact the projected end to end losses will be addressed. This will be updated beginning with 1Q18 MIS. Further, in Method A Model re-development, CAMU is testing an alternative modeling approach to simplify the two-stage severity model structure, and also testing as part of Method C Model development. This will include items 1 and 2. Method A MDD submissions are targeted to be delivered to MRM by 9/30/18.	The Method A 1st Lien/RM model was overhauled in preparation for the 2019 CCAR submission. Specifically, the CMI/CPB/USRB coping strategy along with output sensitivity was tested and documented in accordance with the Loss Forecasting Model Performance Testing Guidance, as documented within MDD chapters 3, 4, and 6.

Severity – Functional Limitation

Limitation ID	Limitation Description	Required Action Item Description	Compensating Control	2019 CCAR Remediation
19544	**Model Performance** The 2nd lien severity model development data contains limited exposure to periods of strong home price growth and low loss severities.	The Sponsor is required to consider overlays for 2nd lien loans in benign scenarios that would account for over prediction observed in end to end loss testing.	Instances where over prediction was observed in end to end loss testing was addressed and independently evaluated based on the Overlay document submitted for the 1st Lien/RM portfolio. Details can be found in the 2018 CCAR Overlay document housed on the Independent Risk site here: https://globalconsumer.collaborationtools. consumer.Bankgroup.net/sites/CCARGCRM/ 2018CCAR/SitePages/Home.aspx	For the CMI portfolio, there are no breaches in end to end NCL/GCL backtesting for short, medium, and long term testing horizons. The model does slightly under predict during the stress period snapshots, and this will be evaluated as part of the overlay process. For the CPB and USRB portfolios, there are some instances of NCL/GCL breaches in the short, medium, and long term testing horizons. In the majority of cases the model is over predicting, and this will be evaluated as part of the overlay process.
19545	**Model Monitoring** The current model testing and monitoring does not provide a complete view of how the model errors impact the projected end to end losses.	Update monitoring to include metrics that provide a more complete view of how model errors impact the projected losses, such the nested error with the DVM model and end to end losses incorporating the partial charge off logic.	Independent Risk is setting up a working group to address ongoing model monitoring, of which a complete view of how the model errors impact the projected end to end losses will be addressed. This will be updated beginning with 1Q18 MIS. Further, in Method A Model re-development, CAMU is testing an alternative modeling approach to simplify the two-stage severity model structure, and also testing as part of Method C Model development. Method A MDD submissions are targeted to be delivered to MRM by	A new distressed value (DV) haircut and partial charge-off logic has been developed and proposed for the Method A model suite. The aforementioned logic is explained in detail within the 1st Lien/RM MDD, which includes nested error as well as end to end outcomes analysis as required by MRM's Loss Forecasting Model Performance Testing Guidance.

			9/30/18.	
19546	**Modeling Approach** Future development should consider expanding the development data to include pre-2008 data and more recent data. More detailed model specifications should be considered for 2nd lien and VA models, and defaults with zero loss should be incorporated into the analysis.	The Model Sponsor is required to consider the points during future redevelopment.	Pre 2008 and more recent data is being included in the re-development of Method A model suite. This will also include a detailed segmentation analysis which will be included in the revised Method A Severity Model MDD which is targeted for delivery to MRM by 9/30/18.	The Method A FHREL model included an expanded data set which is documented extensively in MDD chapters 3 and 4 of the MDD.

Moving on from the prior limitations, as outlined above this model can be used to forecast credit performances delinquency and losses as part of the model's CCAR usage, for mortgage loan populations across various segmentations, vintages and even at loan level. Given the scope and the standard modeling conventions, this model can be used to risk rank individual loans or a specific cohort of loans based on their origination characteristics', collateral value, credit performance and the state of the underlying macroeconomic indicators. This functionality of the model suite can aid with the model's proposed non-CCAR usages which correspond at a broader level to the mitigation of the overall credit risk and improve portfolio performance across existing products and/or new product offerings – in order to execute a disciplined approach towards portfolio growth utilizing the following channels-

- Portfolio Management- Portfolios exhibiting similar loss profiles can be segmented together and appropriate business decisions (asset sale, loan modifications, securitizations, loan loss reserving, troubled debt financings, repurchases risk estimation etc.) can be taken to minimize risks and maximize returns.
- Asset Management –operationalize decisions (people, processes, systems) affecting loss distributions
- Credit Enhancement– manage investor expectations through optimal portfolio management which can improve the debt/credit worthiness of the bank
- Diversification- holistically manage the portfolio size, mix and geographic concentration based its underlying risk/return profile and the policy limits as set by the business

The model can also be used for managing the risk appetite of the mortgage business as set forth in Bank's Global Consumer Credit Fraud Risk Policy (GCCFRP) that in turn can help maintain a prudent capital allocation process through process improvements initiatives or exploring new venues of growth.

While the results of the model can be used for many varied purposes, there are some important limitations where the use of this model is not justified.

a) In particular, this model is solely designed to be used by Bank's mortgage businesses, and should not, by any means, used to predict loss or delinquency for other financial products, such as credit cards or retail demand deposit accounts.
   Compensating Control- Model should be used for its stated designated usages as referenced within the Model Usage Grid in Section 1.1 of the MDD.
b) Estimations of mortgage risk models are based on the utilization of the past historical data, which implicitly assumes that there is a stable relationship between mortgage default rate and the key default risk factors. It is assumed that the relationship between default risk and its key determinants in the future mimics their past relationship. In cases, when this relationship breaks down, the model results can be unstable.
   Compensating Control- Models are considered an abstraction of reality. The logical order of any statistical process is to make predictions that can expect exceptional results to occur as if they were average. People are most likely to take action when the variance is at its peak. Then after results become normal, it is believed that the action was the cause of the change when in fact it was not causal. Hence, model results should always be evaluated in the context of changing macro-economic environment, prevalent lending practices and given industry standards.

c) Any potential change in the regulatory policy, compliance measures (endogenous at the company level or exogenous at the broader economy level- like rate hikes, tax reforms, trade policy changes); will have an immediate effect on how a given financial institution manages its businesses, products and services. The current model usage may be somewhat limited given the fact that the specific outcome(s) or nature of complexity for these future policy events /compliance changes cannot be predicted very well in advance.
   Compensating Control- As part of CAMU's commitment to deliver robust models that align with economic intuition and evolving macro-economic trends, steps are always taken to proactively integrate regulatory preparedness and effects of major policy overhaul or strategic business initiatives as some of the critical components at the beginning of the modeling stage and throughout the model lifecycle. In particular, the model suite are always built with the added flexibility to take into account the effect of overlays as deemed necessary by the respective businesses.

d) Given the fact that the scope of the model extends to Bank Private Bank (CPB) portfolio, there are some inherent limitations involved around the utilization of the CPB data for model development. The CPB portfolio typically caters to ultra-net worth clients who traditionally have much larger balances with very low probabilities of default and losses. As such, the CPB portfolio failed to demonstrate sufficient sensitivities to the historical stress (2007-2008 Financial Crisis) leading to an extremely sparse loss and default data over the entire course of the model development period (2005-2017). The negligible modeling data or lack of thereof violated the Gating Principles, which created impediments around the best way to model this niche population.

Compensating Control- It is very important to note here that while this data-specific limitation is beyond the purview of model development and would continue to exist even in the near future, CAMU paid heed to the concerns raised by model reviewers' on the prior model suites and have developed a systematic approach to navigate around this data limitation. To add some context, CAMU implemented suitable coping strategy for CPB portfolio either in terms of specific interaction variables to model the differentiated sensitivities across certain risk drivers and/or use of splines/caps/floors, where appropriate.

e) Amortization of a loan impacts how its credit risk is calculated for the forecast horizon. For accurate estimation of losses and associated credit worthiness of the modeled portfolios, it is imperative to have adequate and precise logics build around the amortization of real estate loans (Residential Mortgage + Home Equity).

Compensating Control- Based on the production code review results that was shared with CAMU in Q12018, the model development team have already made subsequent changes within the production code to enhance the amortization logic to incorporate state-specific amortization policy, remediate pre-existing discrepancies between how the amortization calculations work within the transition framework (especially for the recovering loans) and how the balances are estimated for future events (delinquency, payoff and loss amounts).

f) The non-modeled transitions pertain to those transitions which cannot be modeled using a regression equation due to limited data availability. On top of the challenges associate with sparse data availability, the data is deemed as extremely volatile. These data related idiosyncrasies created the following modeling challenges, when initial attempts were made to build a statistical model for these non-modeled transitions –

1. Lack of economic explanation or risk drivers that are not correlated with the prevalent macro-economic factors trend in an intuitive way
2. Limited volume, resulting in less than 10 loans (average) transitioning/ month for some of the rare transitions
3. The number of loans transitioning contributes to less than one percent of the total volume(both the source and destination delinquency buckets)
4. Insignificant C- statistics(evaluated based on MRM's MTG criteria)

Compensating Control- Given the afore-mentioned data and modeling challenges based on statistical testing, the alternatives considered (for a detailed explanation on the alternative approaches considered, please refer to Chapter 3 of the MDD) and MRM's feedback on prior year's constant rate approach, an empirically derived waterfall approach was considered for distinguishing between modeled vs non-modeled transitions. Consequently an empirical lookup logic was created for the non-modeled transitions based on the in-sample model development data which have been made to differ across stress and non – stress periods, and also reflected the required granularity & performance differentiation across the following segments (Conventional, GOV, In-Trial). For more information on the non-modeled vs modeled transitions approach, please refer to Chapter 3 and 6 of the MDD.

g) The developed model suite would be benchmarked to an independently developed model suite, which utilizes a different framework to arrive at the same set of outputs/results. This benchmarking helps to establish the rigor and robustness of the model's performance across all-important dimensions.

Compensating Control- The Method A RM model suite will be benchmarked to the newly developed Method C model suite which would use a Hazard modeling methodology to model losses. The Hazard modeling framework differs completely from the transition framework, in that it models the probability of a loan reaching terminal status, given the historical credit performance and other loan level attributes. All stipulated model usages, would be benchmarked to the new Method C model suite, based on the Model Testing Guidance Report, as issued by MRM. Please note, the benchmarking analysis would be completed post the successful validation of the Method C model suite in the first quarter of 2019, as mutually agreed upon by CAMU and MRM.

i. The interest rate cycle is closely related to the economic cycle. In theory, if the economy is growing strongly and inflationary pressures increasing – Central Banks will increase interest rates(primarily the Feds Funds rate[1]) to slow down the economy and prevent inflation. Similarly, if the economy enters into recession with falling inflation and rising unemployment – Central Banks will cut interest rates to provide an economic stimulus to try and increase the rate of economic growth. In 2005 and 2006, the economy was growing strongly and the Federal Reserved increased rates in response to the rapidly growing economy fueled by the surge in real estate lending and corresponding housing prices. With the onset of the 2008/09 recession, the Federal Reserve responded by cutting interest rates close to 0% for an extended time period that ended in late 2015. Starting quarter four of 2015, the Federal Reserve had started to increase interest rates as the US economy returned close to a more 'normal' situation with declining unemployment, increasing GDP and stronger macro-economic environment that led to consistently increasing housing prices.

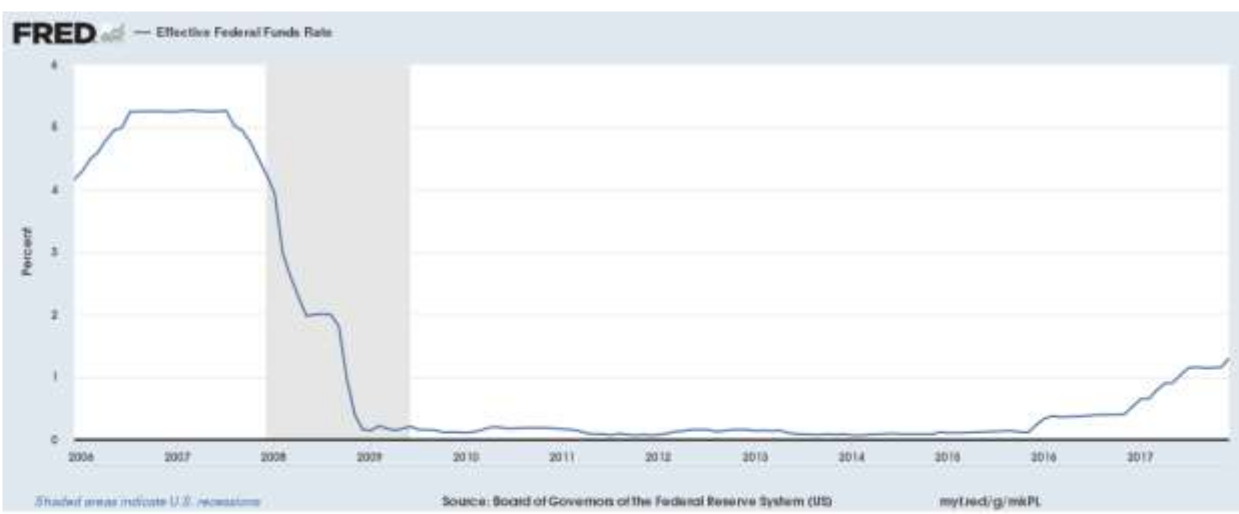
---

[1] The federal funds rate is the rate at which depository institutions (banks) lend reserve balances to other banks on an overnight basis. Reserves are excess balances held at the Federal Reserve to maintain reserve requirements.

The fed funds rate is one of the most important interest rates in the U.S. economy since it affects monetary and financial conditions, which in turn have a bearing on critical aspects of the broad economy including employment, growth, and inflation. The Federal Open Market Committee (FOMC) meets eight times a year, to set the fed funds rate, and uses open market operations to influence the supply of money to meet the target rate.

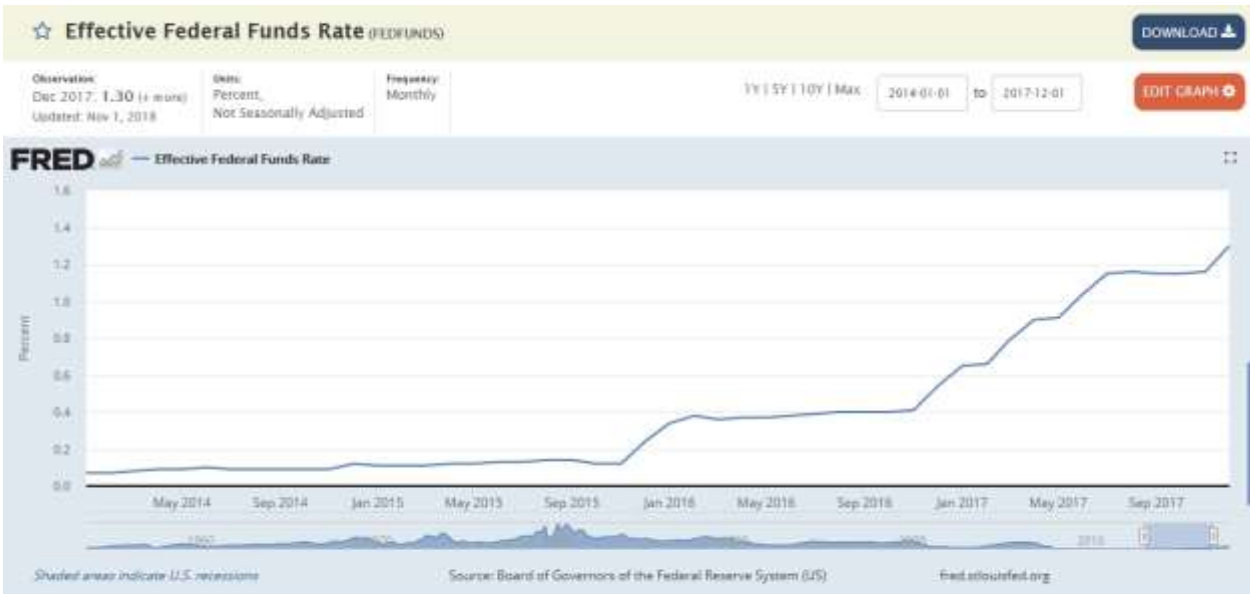
Compensating Control- CAMU is very well aware of the impact of the changing interest rates on the portfolio origination mix (rising rates adversely affect products which have a variable rate component) and the subsequent losses. In order to capture the effect of rates within model development, CAMU has extended its development period to capture an entire decade of rate movements starting from pre-crisis (Feb -2006), through the crisis (2008-09) and post crisis (2011 and beyond). Please see Figure 2.1 below. As shown below, the Fed Funds rate was at 4.49% in Feb 2006, which is the start of the model development data. The rate peaked August 2006 at 5.25% and stayed flat until July 2007. Beginning August 2007, the rates began to drop steadily to prevent the excessive over-heating of the housing market. The recession started Dec 2007 and lasted approximately eighteen months (until June -2009). The recession witnessed a period of significant rate drop as a strategic response to the ongoing housing crisis. Once the crisis ended, rates stayed flat until quarter three of 2015(0.25%). Q4 2015 witnessed the first rate increase post crisis (0.5%). Starting Q4 2015,there has been a steady increase in interest rates[0.75% - Dec16, 1.0% - March 17, 1.25% - June 17, 1.5% - Dec 17]. The model's OOT period, which extends from Jan 2012 to Mar 2014 doesn't impact the model's ability as the rates stayed flat during this time frame. As such CAMU believes that the model development data that extends from Feb 2006 - Dec 2011, and Apr 2014 - Dec 2017 sufficiently captures the rising rates and a complete interest rate cycle.

Figure 2.1: Interest Rate Movement (Source – fred.stlouisfed.org)



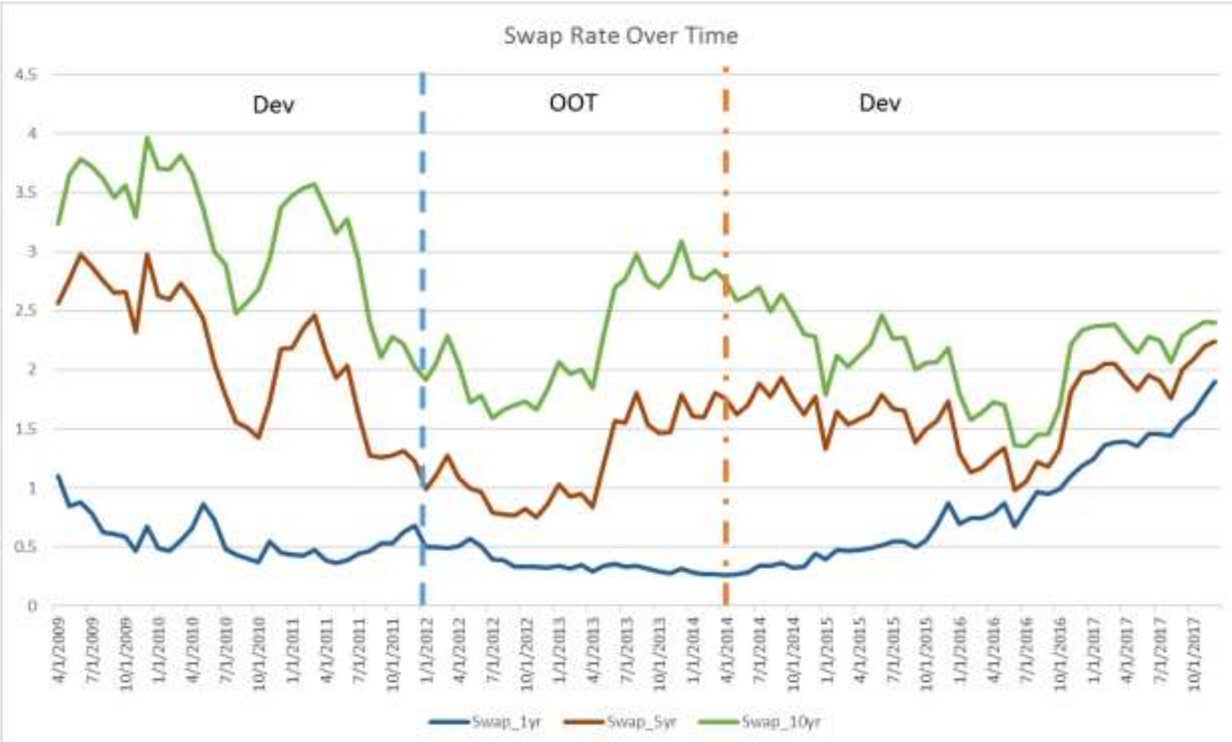
The development data also includes the most recent period of rising rates (2015Q4 -2017Q4). As can be seen in the Figure 2.2 below, the selection of the most recent data period exactly coincides with the recent increases in the interest rates. For additional details, please refer to the following links corresponding to the history of rate increases by the Federal Open Market Committee through its Open Market Operations (OMOs) - https://fred.stlouisfed.org/series/FEDFUNDS and https://www.federalreserve.gov/monetarypolicy/openmarket.htm

Figure 2.2: Recent Period Interest Rate increase (Source – fred.stlouisfed.org)



Several rate variables were used in the model to assess the rate movements in the economy. The market rates used for the interest rate spread calculations as used in the model specifications for the voluntary prepay (VP) models pertained to swap 1 year, 5 year and 10-year rates. The swap curve is a graph of fixed coupon rates of market-quoted interest rate swaps across different maturities in time. The main rationale for using swap rates in the credit risk modeling is because they factor in the counterparty risk that corresponds to the risk that the counterparty to a swap will default and be unable to meet its obligations under the terms of the swap agreement. If the holder of the floating rate is unable to make payments under the swap agreement, the holder of the fixed rate has credit exposure to changes in the interest rate agreement. The spreads calculated in the model are the difference of the swap rate and loan note rate of a similar maturity. The spread captures the interest rate burnout effect in the VP models. The Libor 1 year rate was used part of model implementation to calculate the current note rate for the ARM loans. The London Interbank Offered Rate (LIBOR) is an interest rate based on the average interest rates at which a large number of international banks in London lend money to one another. The LIBOR is among the most common of benchmark interest rate indexes used to calculate the note rate on the adjustable rate mortgages. For additional details on the rate variables as used in the modeled equations, refer to the model specification listing in Section 5.1.4. Please refer to the production code for details around the implementation of the rate variables in the model. Presented below is tabular view of all interest rate variables used within the end-to end RM model. All these underlying interest rate variables and the calculated attribute [P_M_Curr_Note_Rate, P_M_PresIntSpread] aid in assessing effect of rising rates on model's performance. Please see Section

5.1.4[Variable Selection] for additional details. Also illustrated below is the swap rate curve which again reiterates the fact that rate attributes chosen in the RM model align with the interest rate movements in the economy.



#	Interest Rate Variable	RM end-to-end (Brief description)
1	Libor_1yr	Underlying index rate used to calculate current note rate (P_M_Curr_Note_Rate) for ARM loans
2	Swap_10yr	Interest rate spread (P_M_PresIntSpread = note rate – swap_10yr), and the associated Burn-out effect in VP models, FRM>15yr
3	Swap_1yr	The same as above, ARM

4	**Swap_5yr**	The same as above, FRM<=15yr

3. Justification of Modeling Approach

3.1 Description of Applicable Modeling Approach(es) (MAs)

[The purpose of this section is to provide a description of the MAs considered based on specified functional requirements.]

Provide an overview of the applicable MAs, consistent with industry standards, and compliant with regulatory guidance. (If an alternative approach was used, please describe this approach in in Section 3.2.2). In the overview of each MA please include a description of the MA and a discussion of the advantages and limitations.

*[*Required: For Modeling Approach 1, please provide response here]*

::
::
::

*[*Required: For Modeling Approach 2 (if applicable), please provide response here]*

::
::
::

*[*Required: For Modeling Approach [] (if applicable), please provide response here]*

::
::
::

Please identify the preferred MA assuming there were no constraints (e.g., assuming no data limitations), referencing regulatory guidance and industry literature if available.

*[*Required: Please provide response here]*

Credit risk is a critical area in banking and is of concern to a variety of stakeholders: Institutions, consumers and regulators. Central to credit risk is the default event. The primary example of default is when the borrower fails to meet his/her debt obligation and becomes delinquent/defaults on the loan. In the world of mortgage credit risk modeling, different forms of credit risk include the repayment delinquency in retail loans and the loss severity upon the default.

A) Modeling Approaches Applicable to the Probability of Default (PD) Model

BankMortgage has experimented with different classes of PD models over time, from a simple roll rate model to vintage-based models, to a loan level competing risk transition model to the newly developed Hazard model suite developed within the brand new Method C model suite for the 2019 CCAR process.

Since 2014, BankMortgage has used a loan level transition model in a competing risk framework which has been considered to be the main preferred approach for Method A redevelopment this year.

The Method A RM PD transition framework represents the most granular analog of the roll rate models. The account-level transition matrix as illustrated above can be iterated month over month using an expected probability approach to generate the future probability distribution of loans across payment statuses and terminal outcomes each period.

The addition of account-level risk factors contributes to the model's ability to capture portfolio compositional change. As delinquency profile, FICO distribution, loan age distribution, etc., change over time with underwriting tightening or loosening, runoff, and new bookings, account-level parameters help capture the impact of these changes on portfolio outcomes.

Roll rate models can also be extended to take into account macroeconomic scenarios. This is typically accomplished by estimating one or more transition rates for each segment (e.g., current-to-30dpd) as a function of macroeconomic variables using statistical regression models.

Additionally, the model is built with enough flexibility to take into account the impact of various business and policy events as well. For example, accounting policy specifies different charge-off triggers across different delinquency levels, the impact of asset sale onto portfolio risks, etc.

The Method A RM PD transition model leverages the entire spectrum of information reflecting the granular account level factors to the broader macro-economic factors to parametrize the roll rates via regression analyses, typically in the form of discrete choice multinomial equations conditional on the loan status. These transition rates form a matrix, with each element of the matrix representing the transition from a source status to a destination status over a defined time period (e.g. 1 month, 1 quarter). Such a roll rate model allows for positive transition rates for every logically possible transition (e.g., current-to-30dpd, 30dpd-to-60dpd, 60dpd-to-default, etc.). Within the roll rate transition matrix, transitions can be classified as rare and non-rare transitions.

The rare transitions are typically characterized as the following –

1) Jumping transitions from current (or low delinquent) to deep delinquent or IVP,
2) Transitions from deep delinquent to VP,
3) Partially cured transitions from deep delinquent to low delinquent without full cure,

with the succeeding distinguishing features –

- Low volume transitions that could impact the model results significantly when actual default rate and corresponding losses are low in today's environment and portfolio mix. The accurate portrayal of these transitions hence plays an important role in model performance especially for the most recent time period.
- Highly volatile over time with some transitions demonstrating significant performance differences across stress and non-stress
- Some of these transitions do not align with the prevailing macro-economic conditions, , with a few transitions reflecting a better performance during stress compared to recent timeframes, creating further challenges in the model development process

All remaining transitions are classified as non-rare transitions. Within the scope of last year's model all rare transitions were essentially non-modeled and used a constant rate assumption over entire historical period without any differentiation between stress and non-stress periods. Two specific concerns were raised by independent model reviewers regarding prior year's non-modeled transitions. These concerns were related to –

- Overly simplified specific constant rate for stress and recent period could lead to significant model errors
- Lack of a systematic process to determine modeled vs. non-modeled transitions

Based on the above cited feedback, an enhanced step-wise logic was implemented this modeling cycle to systematically define the number of modeled vs non-modeled equations adding further clarity and reasoning for the model's stated level of complexity. This approach ensured that CAMU specifically addressed model reviewer's concerns relegated to last year's constant rate approach assumption.

CAMU strategized the following execution logic [outlined below] for not only defining the modeled vs non-modeled equations (thereby justifying the implied complexity of the model) but also to implement a clear transparent approach for estimating the non-modeled transitions.

Segregate between modeled vs non-modeled transitions based on the following waterfall criteria-Identify modeled vs. non-modeled transitions.

A. Not aligned with business expectations. The first rule of the waterfall logic considers commonly observed business practices to filter out intuitively impossible transitions such as the jumping-to-worse transitions. Not only are these transitions rare, but also tend to be uncorrelated with key risk drivers (for e.g.:- BUK1-BUK6 or BUK7).
B. Not correlated with macro-econ trend in an economically, intuitive way. Typically, historical evidence suggests that stress roll to worse rates should be higher than the non-stress roll to worse rates.
C. Limited volume: on average less than 10 loans/month transitioning from the source delinquency bucket to the destination DLQ bucket

D. Limited contribution portion: The number of loans transitioning represent less than one percent (1%) of the total volume for either the source or destination bucket. Any further granularity will lead to an increase in volatility and overfitting of the applicable data.
E. C-statistics < 0.6 (the variable selection criteria as per MRM's model testing guidance). The model type is logistic regression and all pertinent model diagnostics results are shared in Section 6.1.8.

Loans that fail the defined waterfall are identified as non-modeled transitions while loans that meet the waterfall criteria are modeled using separate regression equations. The waterfall logic is executed over the entire development sample pulled data [02/2006 to 12/2017], which includes both development, the in-time hold out validation and out-of-time validation samples. Additionally, all government and in-trial loans in the data are included in the scope to create a holistic, all –inclusive sample, minimizing potential of omitted data biases.

Given this waterfall logic, two changes were observed within the modeling framework. These were -

i. Some equations which were identified as non-modeled transitions in prior years' were now identified as 'to-be' modeled equations. These represent some key 'curing' transitions which eventually led to improvements in the model's overall accuracy and sensitivity
ii. Some equations which were identified as modeled transitions in prior years' were now identified as 'non- modeled' due to lack of data or limited volume/contribution rule. This significantly aided with reducing the model's complexity in an objective way.

The finalized approach used for estimating the newly identified non-modeled transitions is outlined as below-

1. The Stress values are volume weighted mean of transition rates based on data from 200602 to 201112. The justification for doing so is that the US HPI reaches its peak in the second quarter of 2006 and started dropping over time. It reaches the bottom in 2009 and continues staying low until the first quarter of 2012. Instead of using the exact 27mo stress period defined by MRM, the stress value selection covered this entire period of HPI decrease to avoid over-fitting the model performance due to data volatility or at odd with MRM's "blind back test" requirement.

2. The non-stress/recent values are volume weighted mean of transition rates based on data from 201404 to 201712. The recent four years performance data well reflected the recent portfolio performance and go-forward portfolio mix while avoiding the data volatility and potential model over fitting using only the most recent two years' performance data. To be consistent with other model components, the same 201201-201403 performance data was hold out as OOT data and excluded from non-model transition assumption creation process.
3. Once the performance period is decided, CAMU considered segmentation impact to the non-modeled transition rates and created the stress/non-stress values separately for distinctive segments based on their empirical performance difference. There are three segments for RM included in trial loans, not in trial government loans and not in trial conventional loans. Delinquency information does not truly reflect borrower's payment behavior when a loan is in trial as the change of payment is not finalized and reflected in Bank's delinquency recording system. Therefore, abnormal delinquency movements are expected when a loan is in trial. CAMU has observed different non-modeled rates between government loans and conventional loans as well especially the increase of jump-to-worse transitions for government loans during the non-stress period. Separating government and conventional loans is a prudent choice to account for their difference over time.
4. Due to the uncertain and volatile nature of non-modeled transitions, the stress values are set to non-stress/recent values if the raw stress values are lower/higher than the non-stress values for to-worse/to-better transitions. This is a conservative treatment affecting limited cases (mostly observed for government loans jump-to-worse transitions) to ensure that the stress values are always worse than or at least the same as the non-stress values.

Please note CAMU had shared this approach with IRMO as part of preliminary model review. Please refer to attachment - '3.2 Discussion regarding IRMO's Preliminary Review shared in NA REL Method A Model Development Memorandum' for evidence.

It is important to note here that during finalization of the preferred approach, several alternatives were tried and tested for estimating the non-modeled transitions. All of these alternatives have been discussed in Section 3.2.2 of the MDD.

Given the description of the current modeling approach, it is relevant to add that in the prior years (2017 & 2018), both Method A (Champion) and Method B (Challenger) model suites had leveraged the transition framework. Even though, while the underlying framework was similar, the key difference between these two suites pertained to the strategic implementation of the transition framework itself: while Method A used the expected probabilistic approach for the implementation of the framework, Method B leveraged the simulation framework.

While both Method A and B model suites demonstrated strong back test performances during stress period as well as recent periods and were also able to generate sufficient forecast sensitivity that met Bank internal thresholds for the 2018 CCAR process, limitations were raised by model reviewers on the homologous modeling approach that was utilized in both suites.

To address the specific limitation on the similarity of model framework, for the 2019 CCAR process, CAMU has developed a completely new framework under the Method C model suite which will rely on the Hazard modeling methodology to serve as the primary alternative and challenger to the Method A model suite. The Method C model suite leverages a conditional survival (hazard) framework that includes time varying macroeconomic effects jointly estimated with loan level risk factors to directly model the terminal event losses.

An important distinction between the survival modeling approach and the state transition modeling approach is that survival models predict terminal event outcomes directly. A function of time since observation called the model baseline accounts for the duration between observation and termination events. The intermediate delinquency progression is therefore modeled implicitly as a pure function of time in the survival modeling approach, and explicitly in state transition modeling approach through the one period transition equations.

Presented below is the comparative overview between the primary modeling approach as leveraged by the Method A PD framework and the alternative approach, as deployed by the Method C suite.

Table 3.1.2: Comparison of Preferred PD Modeling Approach with Alternative Approach

Transition Model (Primary Approach)	Hazard Model (Alternative Approach)	Supporting Narrative
Predicts Intermediate (non-terminal) and terminal delinquency States	Predicts terminal events. The intermediate delinquency progressions are modeled implicitly as a function of time	End users expressed a desire for delinquency forecasts for the non-CCAR module. The Transition method allows CAMU to make delinquency predictions consistent with the default and prepayment forecast as specified within the Model Usage Grid.
The model is run month by	All the variables are treated as	The Transition model takes into

month. Dynamic variables in the month after the observation month are updated conditional on previous months' transition outcome, thus not perfect.	static except the future economic variables.	account the time-varying attributes of borrower specific characteristics that affects the terminal events as well as the intermediate events(e.g.: Mark-to-market LTV, Refresh FICO)
Estimated via Multi-Nominal Logic (Competing Risk)	Estimated via Multi-Nominal logit (Competing Risk)	Both use the same statistical regression technique.
A particular loan-month combination creates a single statistical observation.	First, a particular loan month can generate up to n statistical observations per single snapshot date, where n is performance months. Multiple snapshot dates depending on the temporal sampling frequency will further multiply the number of occurrences.	The super panel framework under the Hazard framework quickly "blows-up" the number of observations, requiring extensive sampling.
Use of status history makes the system path dependent.	Not a transition structure	See discussion above

MRM Question - For LGD model, the Dev. sample is Jan2008-Dec2011 & Apr2014-Dec2017 and OOT is Jan2012-Mar2014. Similar rationale as requested above for the PD model should also be provided for the LGD model.

Answer - The non-modeled transitions are only applicable to PD model. LGD model was developed using data from Jan2008 to Dec2017, while 27mo performance data during Jan2012-Mar2014 was hold out for OOT testing. MRM cited a limitation on the current model for not using most recent data to capture recent portfolio mix, performance behavior and macroeconomic trend. CAMU agrees with MRM that the recent data should be included in the model development sample. Loss data prior to 2008 quality is weak as described in the previous data section. 2008-2010 is CMI portfolio stress period and needed for a robust stress model. These aspects limit the use of two end points of development data for OOT exclusions. Needless to say that OOT validation is a regulatory mandate. Therefore, the choice was always limited to pick up a phase in between. The period selected for OOT is excluded from development while variable stability analysis was performed on the full data and various back tests were performed to show overall satisfactory performance across time.

B) Modeling Approaches Applicable to the Distressed Property Valuation

CAMU in prior years had utilized a fully functional DVM model to estimate future distressed property valuation(s) (DVs) for:

1) Loans with an existing distressed valuation on the underlying collateral (referred as the PDV=prior distressed value segment),
2) Loans with no existing distressed valuation (referred to as the NPDV = no prior distressed value segment), at the time of the forecast

This DVM model was used in combination with PD, EAD and LGD models to predict the losses for both CCAR and Non-CCAR usages of the model. While the prior year's model had demonstrated reasonable back testing performance and sufficient sensitivity in stress scenario forecasts over the past three years, as part of the model validation process, definitive recommendations were made to improve the variable selection process which hinged on a single macro-economic factor – HPI and a time decay factor. Given the feedback received from model reviewers', the DVM model from prior years has been replaced with newly formulated haircut lookup logic. This haircut logic is marked-to-market and has been estimated based on empirical data for both PDV (prior distressed valuation available) and the NPDV (No prior distressed valuation available). This new logic has the flexibility to directly incorporate home price updates at various stages of the forecasting process. The Distressed Valuation haircut (DV) rates are estimated using the same dataset as what is used for current Collateral Risk Management processes (starting from 2010) as well as observations from a legacy data source (DRI, for observations dated 2009 and earlier). The incorporation of both DM and DRI data sources in the estimation of the DV logic was made to align with Collateral Risk Management's valuation practices, who use the DM data for estimating all property values starting 2010. For the stress period, CAMU continued to leverage the DRI dataset for obtaining property values pre-2010. Based on all conceding conversations with the Collateral Risk Management team, it was recommended to utilize the official and validated data source following the collateral waterfall logic. See Attachment – '3.1 FW CRM recommendation for DV Logic' for the official CRM Recommendation document.

CAMU created a comprehensive list of loan attributes/key risk drivers listed as below, which were tested on a stand-alone basis to evaluate its impact on the DV of the property for both the stress and non-stress periods, which use a combination of both the afore-mentioned datasets-

- Lien Position
- Occupancy Type
- Origination Channel
- Property Type
- Loan Balance
- LTV

- Loan Age
- Loan Type (Government versus Conventional)
- PMI Coverage
- Modification Status
- Business Entity (CMI or other)
- State of the property (judicial vs non-judicial)

It was noted that for both stress and non-stress periods), similar sensitivities were noted across most risk drivers. For more information on the haircut logic, the methodology, mathematical notation, please refer to Section 5.1.2 of the MDD.

C) Modeling Approaches Applicable to the Loss Given Default (LGD) Model

CAMU had experimented with different classes of severity models over time. Since 2015, BankMortgage has used a 2 stage loss severity model for its North America Method A model suite. For the first lien mortgage loans, the first stage estimated the probability of loss outcome (involuntary payoff) in a competing risk framework for the following loss disposition types: Real Estate Owned (REO), Charge-off (CO), Short Sale (SS) and Third-party Foreclosure (3F). The second stage then estimated the loss severity for each of these different loss outcomes using non-linear regressions. The total loss estimate was the product of these two stages. For the 2nd lien loans the first stage estimated the probability of full loss and non-full (partial) loss and the 2nd stage estimated non-full loss severity. Zero losses have traditionally been excluded from the scope of the model development data within the existing modeling framework for both first and 2nd lien loans. The motivation for leveraging the two-stage loss modeling approach was not only based on its parallel correspondence with the policy driven mandates undertaken by the NA mortgage business towards recognizing losses but also on the model's factual equivalence to a widely used vendor model developed by Black Knight Financial Services that utilized a similar approach to estimating residential mortgage losses.

Presented below is the diagrammatic illustration of the 2017 Method A LGD modeling framework.

Figure 3.1.1: Summary of 2017 Method A LGD modeling approach

Champion Approach
<u>Severity Estimation:</u> **Two steps to estimate total loss** • 1^{st} Lien - 1^{st} step: estimate prob. of IVP disposition type (logit) - 2^{nd} step: estimate severity by IVP disposition type (non-linear) • 2^{nd} Lien - 1^{st} step: estimate prob. of full loss and non-full loss (logit) - 2^{nd} step: estimate non-full loss amount (non-linear)

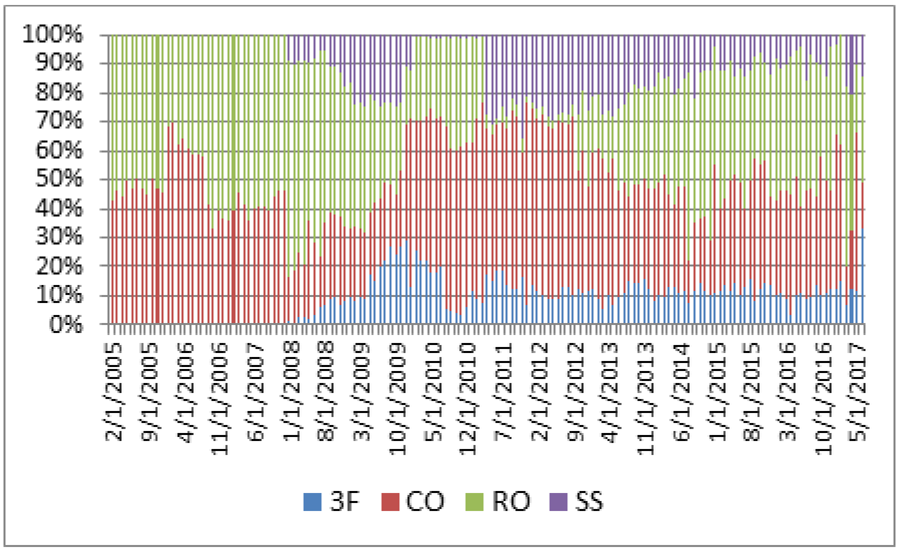
Despite the strong corroboration towards the afore-mentioned approach as evidenced both from the business policy perspective and common used industry practices, in recent times questions and comments were raised both by the model reviewers'/ end users as well as the internal CCAR modeling team as part of the self-improvement initiatives, towards the credibility of this framework (first liens only) in modeling loss disposition types which are guided heavily by the changes in the internal business policy changes and external regulatory landscape.

Taking into account the features of the existing framework and the model reviewer's feedback, the modeling team proactively conducted definitive research and analysis on the varied alternative loss modeling approaches and proposed to use an alternative two-step loss severity rate approach for the first lien loans. The new framework for the first lien loans would continue to leverage the two stage approach towards modeling losses wherein the first stage segments the losses based on their outcome [full, partial and zero losses] and the second stage models the loss severity for the partial losses only. This new framework helped mitigate the shortcomings associated with forecasting the future loss behavior by disposition types, the challenges associated with modeling loans with zero losses and LTV effects.

As mentioned before the build of this framework for first lien loans were primarily motivated by the flaws associated with the existing frameworks which are illustrated as below.

- As mentioned earlier, the loss disposition type is a policy driven mandate which has continued to change and evolve over time as guided by the business's current lending practices, regulatory landscape and the macroeconomic environment. The figure below illustrates the historical trend associated with the different loss disposition types [Third-party Foreclosure (3F), Charge-off (CO), Real Estate Owned (REO), Short Sale (SS)] for the NA Mortgage business. The volatilities across the different loss dispositions over time implies that there is no consistent relation between disposition type and actual losses incurred which in turn implies that a historical model fit would not be able to accurately reflect the policy changes over time. As a result the model forecast may not be stable. Since a statistical model based on historical loss type distribution, especially for the periods prior to 2008 (see Figure 1) which only had two loss disposition types [CO, REO], may not be applicable to future unknown loss mitigation strategy changes, this justified the need to switch from a disposition type modelling to a two-step full/partial/zero loss severity rate type approach for the Method A first Lien loans. Another rationale that helps strengthen the above argument is the fact that the model's proposed business usages doesn't segment by loss disposition type, which completely eliminates the need to model loss outcomes by disposition type separately.

Figure 3.1.2: Historical Trend – Loss Disposition Type (first Lien Only)



LEGEND -
3F= Third Party Foreclosure
CO=charge off
RO=REO (Real Estate Owned)
SS = short sales

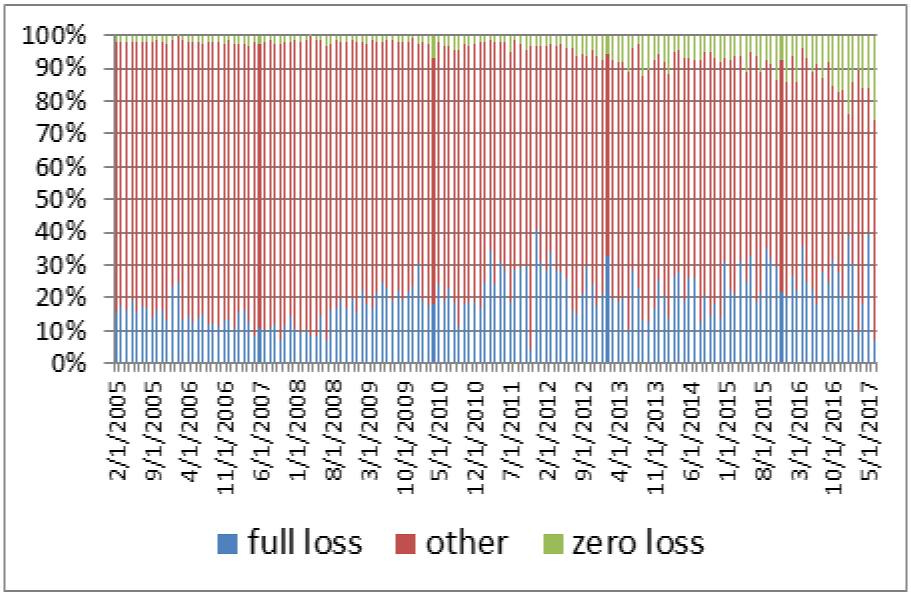
Another data specific challenge with the loss disposition type attribute is that the disposition type is not static over time. As indicated in the table 3.1.3 below, approximately 4.4% of the IVP loans underwent change in disposition type over time, which created challenges using the existing modelling framework. This problem is mitigated using the newly proposed modelling framework which classifies loans based on the Lifetime loss = cumulative losses up to 6-month after IVP date.

Table 3.1.3: Loss Disposition Type Change over Time

flag	Frequency	Percent	Cumulative Frequency	Cumulative Percent
CHANGE	25753	4.41	25753	4.41
SAME	557631	95.59	583384	100.00

- Third, there has been an influx of loans with zero losses in recent times (> 15% of total IVP loans) with the improving macroeconomic environment as evident in the improving HPI in recent times (illustrated in the following figure). In the prior version of the model, loans with zero losses only account for a minimal portion and posed modelling challenges, hence were excluded from the model development data. However, CAMU acknowledged the fact that excluding loans with zero losses is not a prudent modelling choice especially given the recent home price appreciation environment and therefore instilled changes to the modelling framework to accurately model the loans with zero losses.

Figure 3.1.3: Historical Trend – Zero losses (first Lien Only)



- Another segue for the loans with zero losses is that most of these loans (excluding VA loans since they roll under a different accounting rule) belong to the charge-off status. With the existing modelling framework, first lien loans with zero losses were all excluded from the development data as it is considered highly unlikely for loans that have been already charged – off to have zero losses. In Table 3.1.4 below, approximately ~2% of the charged-off loans have been associated with zero losses. While there is no definitive hypothesis or conclusive rationale supporting this data anomaly, it was considered a prudent modelling approach to not exclude these loans as part of the model development data. Using the new modelling framework, all of these loans will be retained as part of the development data pull thus enriching the dataset with increased number of records and also aid in a more robust model development process that aligns with the existing portfolio composition.

Table 3.1.4: Distribution of loans with zero losses

Non VA	Full loss	Partial loss	Zero loss	Total

only				
3F	1,158	8,192	582	9,932
CO	191,635	49,783	4,502	245,920
RO	811	34,293	914	36,018
SS	9,334	32,982	860	43,176
Grand Total	**202,938**	**125,250**	**6,858**	**335,046**

- Using this new modelling framework that projects losses using a two-step full/partial/zero loss severity rate approach, the model's sensitivity and monotonicity has significant improved with respect to one of the key risk drivers of the LGD model framework - aka borrower's LTV. Intuitively speaking, higher LTV is associated with higher severity rate due to relatively higher debt compared to the total property value. The figure below illustrates the bivariate relationship between risk log-odds ratio (loss severity) and current LTV for the existing framework relative to the new framework. Intuitively speaking, periods of rising HPI, indicate improved home valuation which in turn lowers the mark-to-market LTV (current loan balance/home value) and the corresponding loss rates. Under the existing modelling framework which modelled all loans by their disposition type, the empirical data demonstrated a 'V' shaped relationship between the LTV and loss severity. In other words, the left hand side of the figure 3.1.4.1 implied that for all the first Lien loans in the lower LTV bins, as the LTV increased, the loss severity decreased which had a counterintuitive implication given the common business understanding associated with rising HPI. Segregating the severities by each disposition type (Figure 3B) under the existing two-stage framework, continued to reveal a similar counterintuitive relation between borrower's mark-to market LTV and loss rates.

Figure 3.1.4.1: LTV Effect: Existing vs Alternative Framework (first Lien Only)

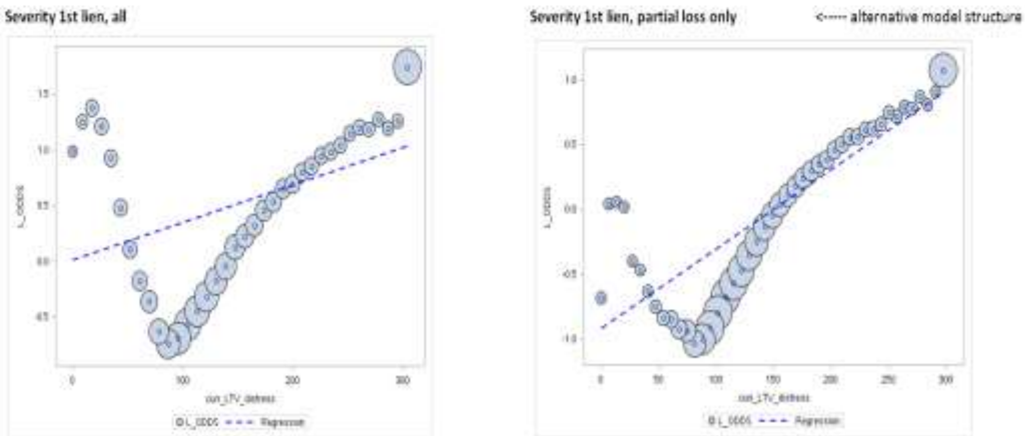
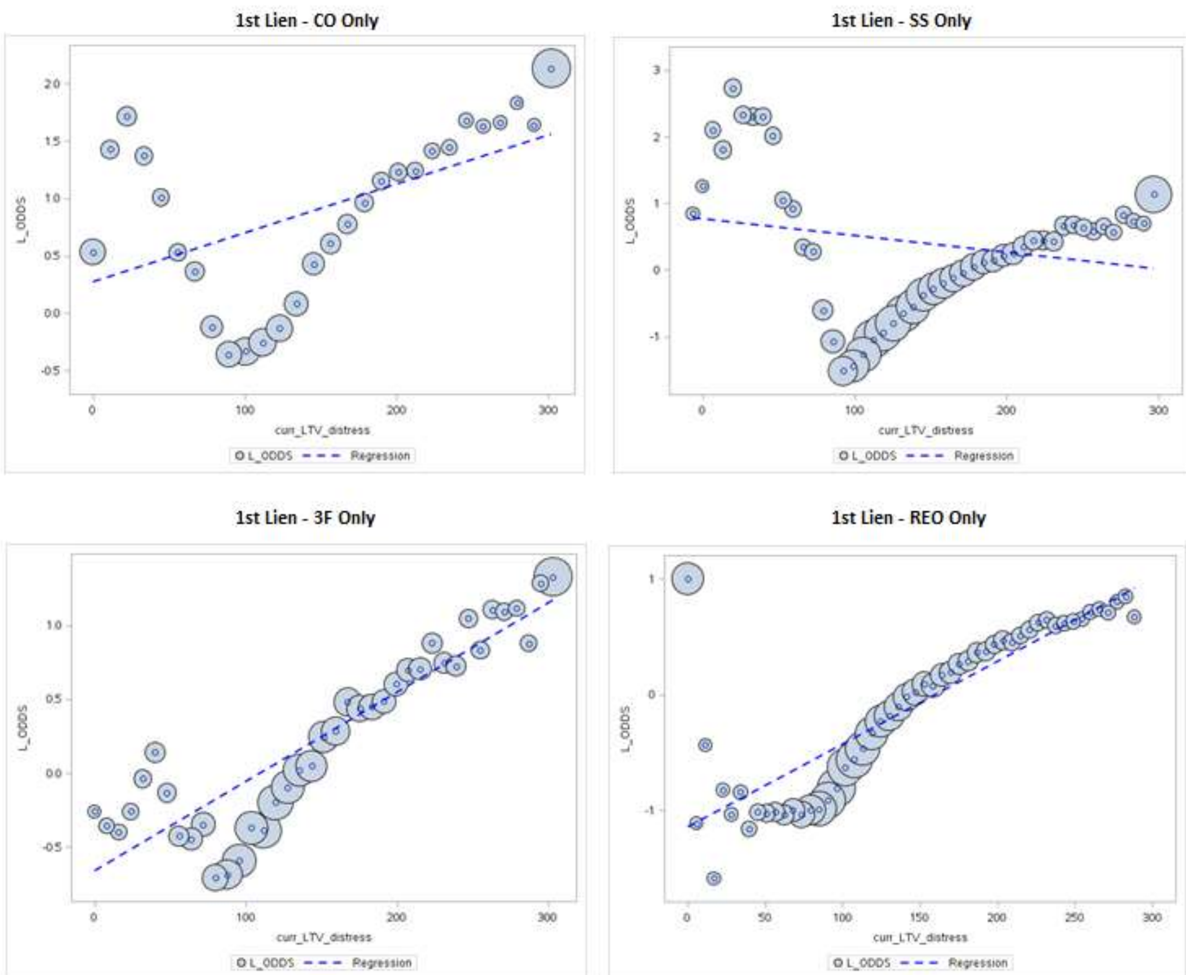
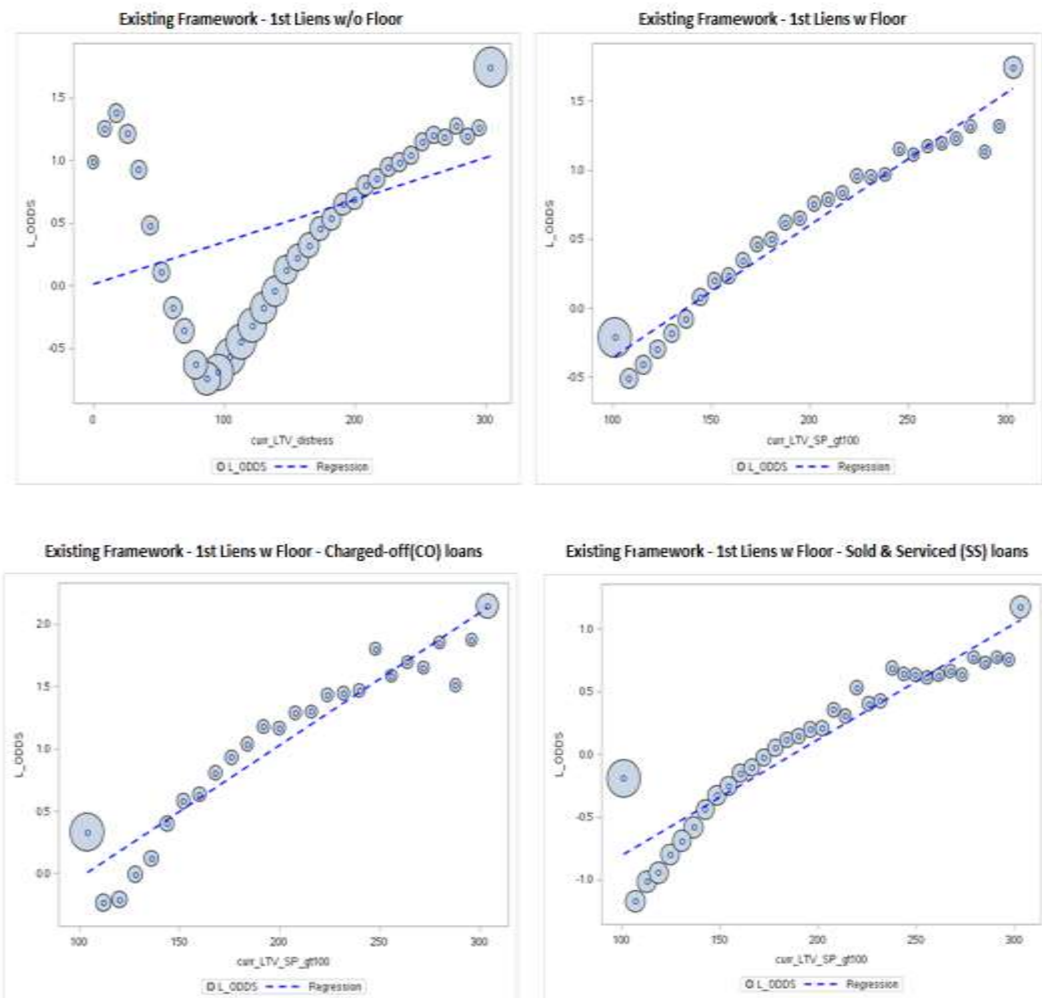


Figure 3.1.4.2: LTV Effect by disposition Type: Existing Framework



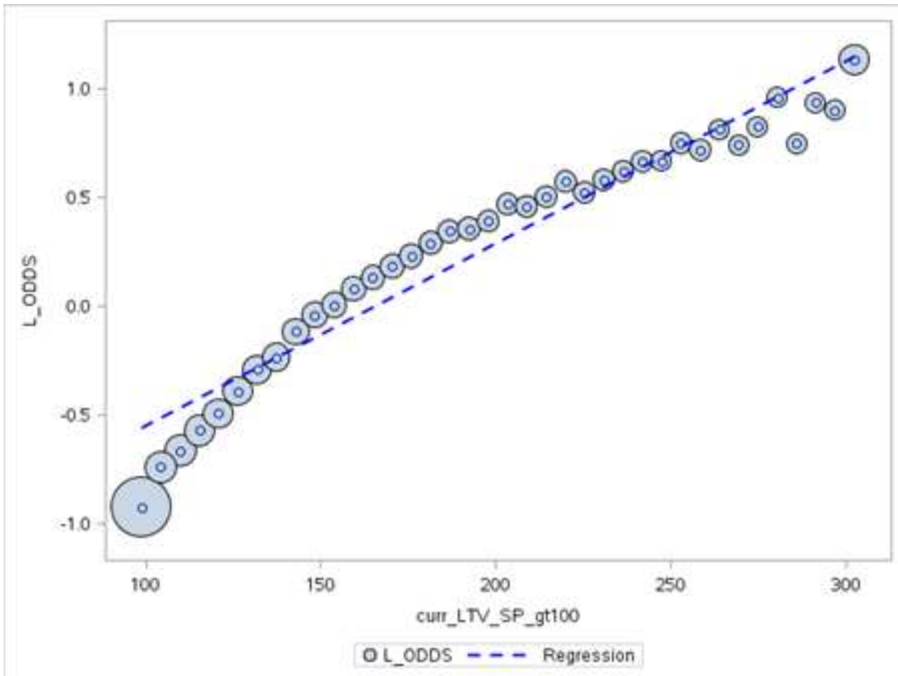
Hence to amend this V-shaped anomaly in the empirical data, it was proposed to create a user defined floor at LTV = 100. However, despite the creation of this floor, the model would continue to render counterintuitive connotation with respect to the LTV variate. This would impact the model's monotonicity (In Figure 4 below loans with LTV < 100 have higher loss rate compared to loans with LTV at 120 specifically for both CO and SS loss disposition types respectively, as illustrated in the bottom two charts for Figure 4) and overall sensitivity with respect to the HPI attribute, especially during periods of improving HPI.

Figure 3.1.4: LTV Effect with and without Floor by Disposition Types: Existing Framework (All first Liens)



Since a definitive solution could not be devised to nullify the V-shaped effect of borrower's LTV to loss rate, it was deemed necessary to alter the modelling framework for the first lien loans to accurately model the LTV effect for the losses (partial) observed. The figure illustrated below provides a clear demonstration of the proposed model's correct rank –ordering (monotonicity) with the floor set at 100 for modelling the partial losses (Full/Zero losses are not considered within the spectrum of the Stage 2 model).

Figure 3.1.6: LTV Effect with Floor under new Framework (first Lien only)



All the above rationale and supporting empirical analysis provided the necessary needs-based evidence to the model developers to effectively switch from the existing approach to the new approach for modelling loss rates for the 2019 Method A CCAR model suite.

3.2 Evolution of Modeling Approaches Attempted

The purpose of this section is to provide an overview of the alternatives considered in arriving at the chosen MA.

3.2.1 For the preferred Modeling Approach identified in Section 3.1, please describe the challenges faced, if any. (If the preferred MA is not your chosen MA, at least one of the following must be addressed in support of the chosen MA.)

This question is not applicable. The approach discussed in the prior section for both PD and LGD models are the preferred modeling approaches.

The Model Testing Guidance which includes the Gating Principles, Model Performance testing and Code and Data Guidance is attached below:

Retail Loss Forecasting

MTG - Retail Loss Forecasting.docx

PPNR

MTG - PPNR.docx

Inputs:

[Model Sponsor must adhere to the Gating Principles guidance and the Gating Principles excels provided in the Model Testing Guidance. In case Coping strategy has been employed to address the lack of portfolio data and/or stress period, Model Sponsor must clearly provide the rationale for employing Coping strategy in terms of risk behavior of customers (for e.g. CLTV etc.), regulatory guidance etc. Model Sponsor must demonstrate the appropriateness of Coping and provide the GP excels for both the coping market and the target market as required by the Guidance. For example, if Australia is coped with Taiwan data, then in this case GP excels for Australia as well as Taiwan must be provided. For detailed information of Gating Principles please refer to the Gating Principle section of the Model Testing Guidance. The Gating Principles templates are attached under section 11.5 of the Model Testing Guidance]

Gating Principle 1: Current Portfolio Profile

- Portfolio must be in a mature state, i.e. should be sufficient in terms of data history, size of the portfolio; history of new and existing accounts, active, pay off, and default accounts etc.
- If the portfolio was not observed to be in mature state in terms of number of defaults/ sufficient amount of balance or revenue, was coping strategy employed?
  - In case of account level models; If portfolio is viewed as not having sufficient number of target variable population (e.g. number of defaults, etc.) the development data must be augmented
  - In case of portfolio level models; If portfolio is viewed as not having sufficient data points, the development data must be augmented

*[*Required: Please provide response here]*

The Residential Mortgage (non-Mod) population is mature and has stable history of active accounts, defaults and payoffs covering the period of Feb 2006 to Dec 2017. This includes a complete business cycle of the mortgage industry including the pre-financial crisis and post financial crisis events.

Table 3.2.1.1 Model Development Data Usage for Residential Mortgage

Model	Data use	Observation window	Gating Principle
RM PD Model	Model development In time validation Out-of-time validation	Feb 2006-Dec 2011 & Apr 2014 - Dec 2017 (80%) Feb 2006-Dec 2011 & Apr 2014 - Dec 2017 (20%) Jan 2012 - Mar 2014	Feb 2006 - Dec 2017
Severity Model	Model development In time validation Out-of-time validation	Jan 2008 - Dec 2011 & Apr 2014 - Jun 2017 (80%) Jan 2008 - Dec 2011 & Apr 2014 - Jun 2017 (20%) Jan 2012 - Mar 2014	Jan 2008 –Jun 2017

Table 3.2.1.2 Development Data Time Period for each Model Component

Model	Development Period
Probability of Default Model	Feb 2006 – December 2017
Loss given Default Model	Jan -2008 - June 2017
DV Logic	Stress Period -Jan 2005 – December 2009[DRI data source] Recent(Non-Stress) – Jan 2010 - December 2017[DM data source]

For the purpose of the Gating Principle template, the "Model Development Data" refers to all the periods used in developing the models including: In-time validation period, Out-of-time validation period and model development period. Table 3.2.1 represents the model development data usage periods for both PD and Severity model.

Figure 3.2.1 Residential Mortgage Development Data Default Rate and Unemployment Trend

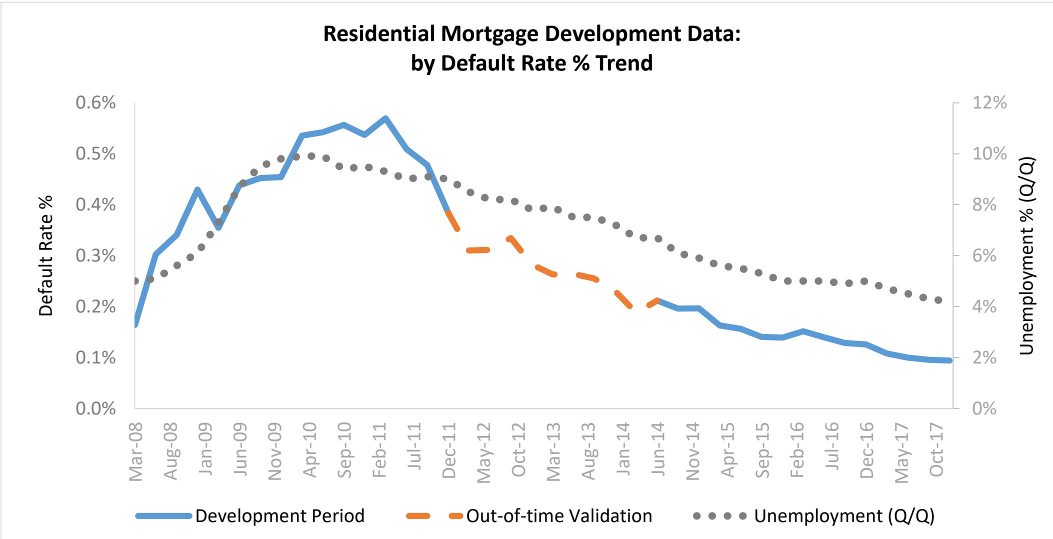


Figure 3.2.1 represents the development data period used in the NAM Residential Mortgage Method A both the development and out-of-time validation period. As depicted in the chart above, it represents a complete spectrum of the mortgage business cycle, specifically, during the US economy stress period during the recent financial crisis experienced between 2008 until 2010. Included in the chart above is the US unemployment rate (Q/Q change) as point of reference and the historical default rate trend used in model development. The reason for selecting the development period was based on the following rational:

1) It enables the model to be developed on a wider spectrum of economic conditions and loan origination credit policy.

2) It includes more loans originated after the crisis so that the model development data can better resembles the go-forward portfolio as of today.

3) It utilizes more loans that have gone through the ARM/IO reset process so the model can better capture the reset risk associated with the portfolio

Detailed results and explanation are also described in section 6.2.2 of this document.

Table 3.2.1.3 Residential by Business Portfolio (CMI and CPB)

By Business Portfolio	***ENR $MM***									
	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
CitiMortgage (CMI)	104,160.0	85,616.9	72,632.7	69,432.2	65,745.5	54,738.7	51,244.6	51,949.0	50,776.1	47,655.5
Citi Private Bank (CPB)	14,239.0	12,109.7	11,550.0	12,431.6	14,578.6	16,432.0	17,910.2	19,993.5	22,913.8	25,352.6

	30+ Delinquency $MM									
	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
CitiMortgage (CMI)	11,070	15,234	9,901	7,669	6,198	4,249	2,980	2,144	1,483	1,211
Citi Private Bank (CPB)	229.5	277.1	345.2	333.4	260.6	227.9	164.8	158.2	158.0	184.1

	NCL $MM									
	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
CitiMortgage (CMI)	1,176.5	2,364.6	1,250.1	675.7	664.6	314.8	131.1	49.5	29.6	21.2
Citi Private Bank (CPB)	0.3	2.3	4.3	7.7	17.1	16.2	6.0	5.8	1.7	1.8

Source: 3.2.1 Gating Principle 1&2 Method A - CitiMortgage (CMI), Citi Private Bank(CPB)

The NAM Residential Mortgage Method A combines all main business portfolio namely: Bankmortgage (CMI) and Bank Private Bank(CPB) in the model development process. Although the model aggregated all the portfolio segments, individual business portfolio segments were reviewed diligently against the Gating Principle requirements to make sure each portfolio is mature, have sufficient data in terms of historical performance, sufficient history of new and existing accounts, active, pay off, and default accounts etc. Table 3.2.1.3 is an excerpt from individual Gating Principle template 1 & 2 for each portfolio segments: CMI and CPB. The table above shows historical performance (delinquency and loss) availability and individual portfolio size over the same time period used in model development, showing both in stress and non-stress periods. For more detail information regarding the individual portfolio segments, including detail historical performance, information on new accounts, portfolio information (Existing Active accounts, ENR, Delinquency information, GCL, Recovery and NCL) and individual portfolio strategic plan available, refer to the Attachment 3.2.1 Gating Principle 1&2 Method A – BankMortgage(CMI) – Residential Mortgage, Bank Private Bank(CPB) – Residential Mortgage.xls.

Each model component is reviewed and examined carefully in the development period to ensure there is sufficient data for performance. As showed in Table 3.2.1.3, both MOD PD model and LGD model satisfied the minimum sample size requirements in Gating Principle 1. The aggregated GCL in terms of unit and of amount in Table 3.2.1.4 provide further support that the data follows the economic cycle and is sufficient for modeling. The DV Logic data captures the log ratio of end appraisal value and beginning appraisal value at stress and non-stress periods. As illustrated in Table 3.2.1.5 and Figure 3.2.1.2, the data is also sufficient for modeling and follows the economic cycle.

Table 3.2.1.3 Minimum Requirement in Gating Principle 1 for PD and LGD Model

Area	Field	Guidance and examples	Response
Data identification	**Period of model development**	*Example: Jan 2008 - Mar 2016*	Jan 2005 - Dec 2017
	Number of Observations for Development*	*Example: 234567*	11,187,779
	Number of defaults/bads or any other target variable*	*Example: 891*	171,631
	Portfolio Event Rate	*Calculated*	1.53%
	Minimum number of Observations for Development	*Calculated*	48,528
	Minimum Number of defaults/bads or any other target variable	*Calculated*	745

Sufficient Observations
Sufficient Target Population

Note:
(1) Number of Observations for Development is the sample down samples from Feb 2006 to Dec 2017.
(2) Number of defaults/bads is the total defaults in the first mortgage (CPB and non-CPB)

Severity LGD Model: entire samples for MOD and non-MOD populations

Area	Field	Guidance and examples	Response
Data identification	**Period of model development**	*Example: Jan 2008 - Mar 2016*	Jan 2008 - Jun 2017
	Number of Observations for Development	*Example: 234567*	304,926
	Number of full loss	*Example: 891*	187,200
	Portfolio Event Rate	*Calculated*	61.39%
	Minimum number of Observations for Development	*Calculated*	117,712
	Minimum Number of defaults/bads or any other target variable	*Calculated*	72,266

Sufficient Observations
Sufficient Target Population

*Full loss (loss_ind = full)

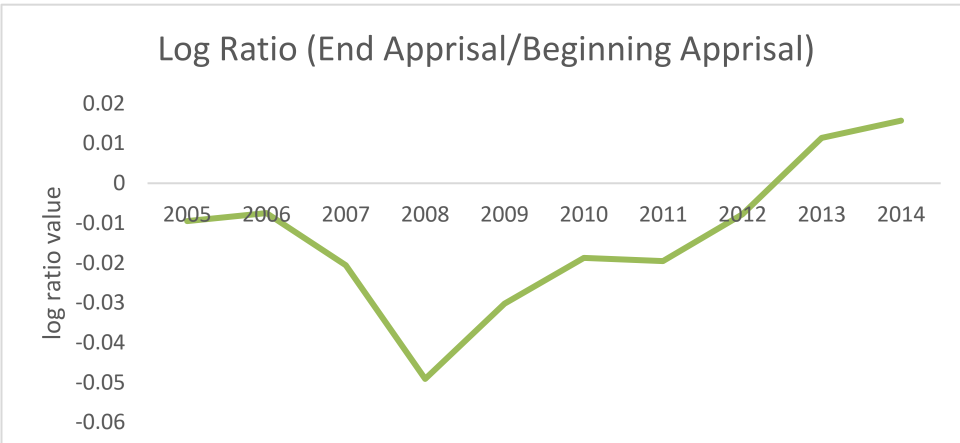
Table 3.2.1.4 Loss in LGD Model

yr	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Total Loss Unit	46,912	71,768	59,719	46,192	31,714	18,935	11,573	7,815	5,165	2,204
Total Loss Amount	3,248,450,999	5,358,111,961	4,202,218,976	3,242,233,454	2,108,485,234	1,106,032,547	611,629,454	401,894,423	262,711,973	107,659,291
Principal Balance at Year End	$590,002,933	$581,613,580	$431,976,162	$324,607,427	$179,237,895	$93,433,646	$75,087,522	$41,650,502	$23,619,706	$32,497,325

Table 3.2.1.5 Appraisal Values in DVM Model

Yr	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Total Number of Observations	20,696	34,111	43,587	64,676	127,565	186,537	191,450	180,187	147,047	53,379
early appraisal value	102,205	117,782	154,076	183,948	185,752	189,914	186,725	186,415	184,056	198,683
appraisal value at year	100,002	115,766	146,955	164,278	173,258	181,903	178,512	183,079	188,961	206,005
dependent variable: ratio in log	-0.0095	-0.0075	-0.0206	-0.0491	-0.0302	-0.0187	-0.0195	-0.0078	0.0114	0.0157

Figure 3.2.1.2 Log Ratio of Appraisal values



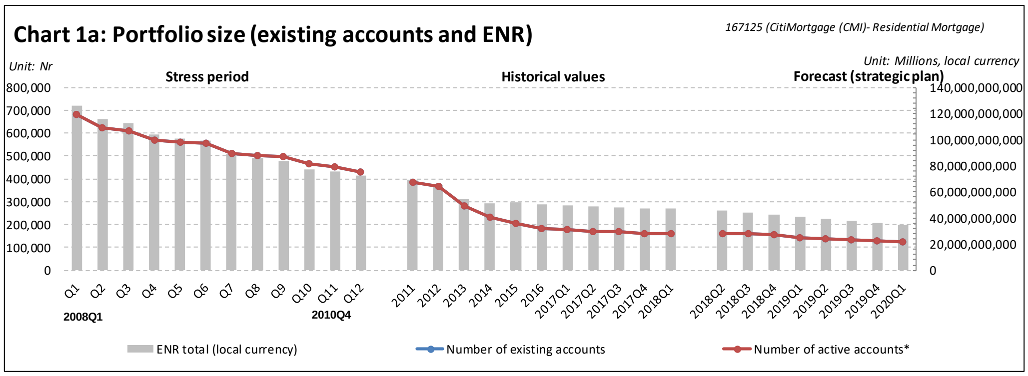
Gating Principle 2: Future Portfolio Profile

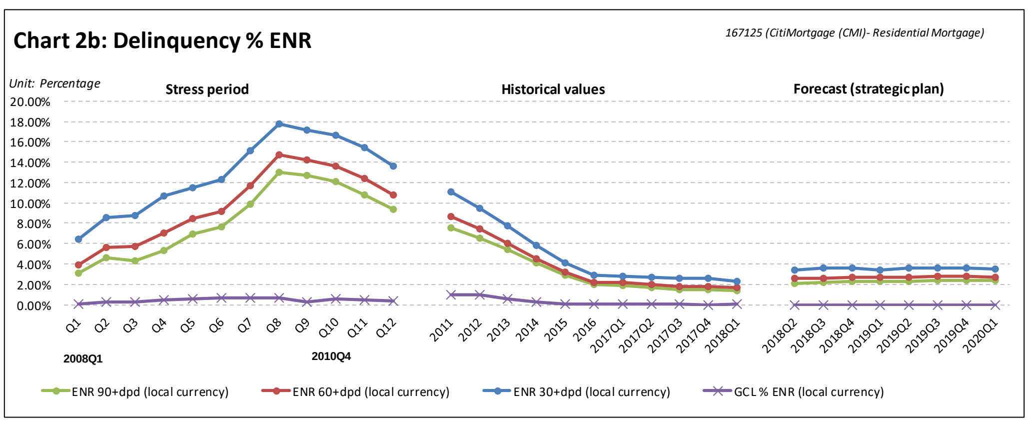
- Please describe whether or not the "go-forward" portfolio set is similar to the development portfolio. Please consider a comprehensive set of potential business changes that could impact the MA including changes in portfolio characteristics (e.g., average limit, exposure, remaining maturity, interest rate, LTV, FICO, delinquency, CD maturity, time-on-file), product mix, product features (e.g., terms and conditions), underwriting criteria, regulation, business / collection / risk management strategy, marketing campaigns, macroeconomic / industry / political conditions, etc.

*[*Required: Please provide response here]*

The development data set includes a complete business cycle including both stress and non-stress periods over the past eleven years. The development data also includes recent loan performance data through 4Q17, which sufficiently represents all historical changes in credit policy as well as enhancements to business strategies over time (i.e. business acquisition and asset disposition). Based on the business strategic plan presented by the respective loss forecasting teams, the overall development sample represents the most recent portfolio trend and the stress scenario based on the most recent portfolio characteristics, underwriting criteria, regulatory environment, macroeconomic trend and risk management strategy. Regarding business strategies employed to date, the residential mortgage portfolio has reduced significantly in the last few years. The go-forward state of the portfolio is representative of the development sample as it includes loan performance up to time of asset sale / modification, as well as recent loan performance. Underwriting criteria and credit policy have largely tightened post financial crisis, however, have remained stable in the prior five years. The future planned originations are projected to slightly increase through 2020, but are not near the record highs we have seen in previous years. Finally, the macroeconomic trends and sensitivities are discussed in further detail in MDD section 6.4.

Table 3.2.1.2 Gating Principle 1&2 Portfolio Size and Delinquency % ENR Charts





As stated in GP1&2 and the rationale behind the selection of the development sample, one of the main reasons of choosing performance that is more recent is that it will resemble more current and the go-forward portfolio. As depicted in table 3.2.1.2 Model sponsor strategic plan resembles most recent performance (in terms portfolio size ENR and delinquency performance) including different changes in current policy and business strategy. On the other hand, including data from the stress period will provide some conservative cushion and sensitivity to potential crisis or stress happened in the future.

Although the portfolio experienced a dramatic shift in performance and portfolio size in previous 10 years mainly due to a different regulatory environment, underwriting policy and difference in risk management strategy, this has changed significantly resulting to smaller volume but better quality loans reflecting current and most recent "go-forward" state of the portfolio with no major changes in the current policy or external regulatory environment. For more historical details and go-forward available portfolio strategic plan, including detail portfolio characteristics (Avg LTV, Avg FICO, etc.) refer to the Attachment 3.2.1 Gating Principle 1&2 Method A – BankMortgage(CMI) – Residential Mortgage and , Bank Private Bank(CPB) – Residential Mortgage.xls.

Gating Principle 3: Stress Period Identification (Applicable for Stress Models only)

- The portfolio dataset must have performance history through at least one adverse macro-economic stress period as defined by MRM for each market. Developer/Sponsor must refer and use this stress period. Please refer to Gating Principles section of the Model Testing Guidance for stress period identification

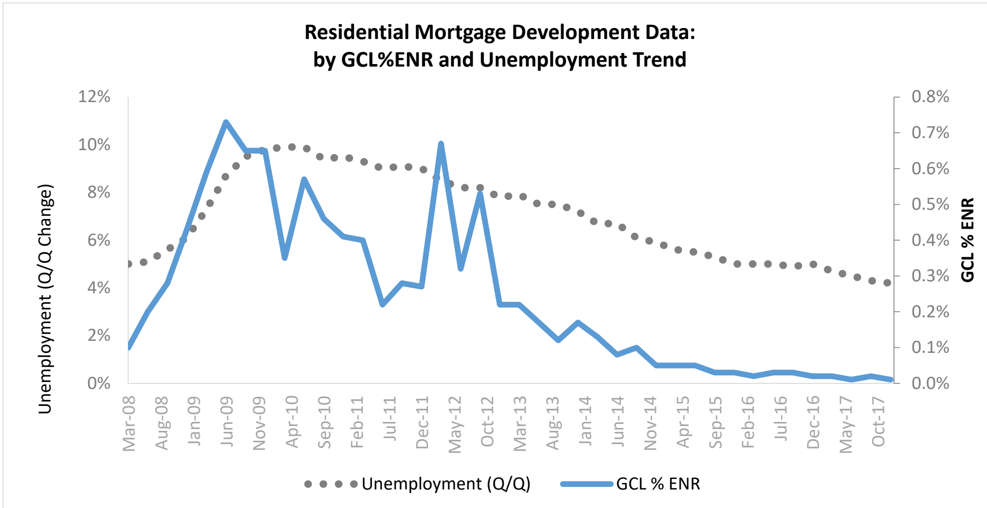
- If there is no macro-economic stress period identified in the portfolio data, coping strategy must be considered.

- *If the stress period for a market is not available in the MRM defined stress period table or if the developer disagrees with MRM defined stress period; then the developer must provide the evidence of stress period being considered for their portfolio. Same should also be supported with data analysis.*
- The portfolio dataset must have performance history through at least one adverse macro-economic stress period. A stress period is defined as one that has:
  - At least two consecutive quarters of negative GDP growth rate, OR
  - At least two quarters of GDP growth rate that falls below one standard deviation from its historical mean (10 years).
- If macro-economic indicators other than GDP are believed to represent a stress period in a particular economy, evidence must be provided to demonstrate the presence of stress.

*[*Required: Please provide response here]*

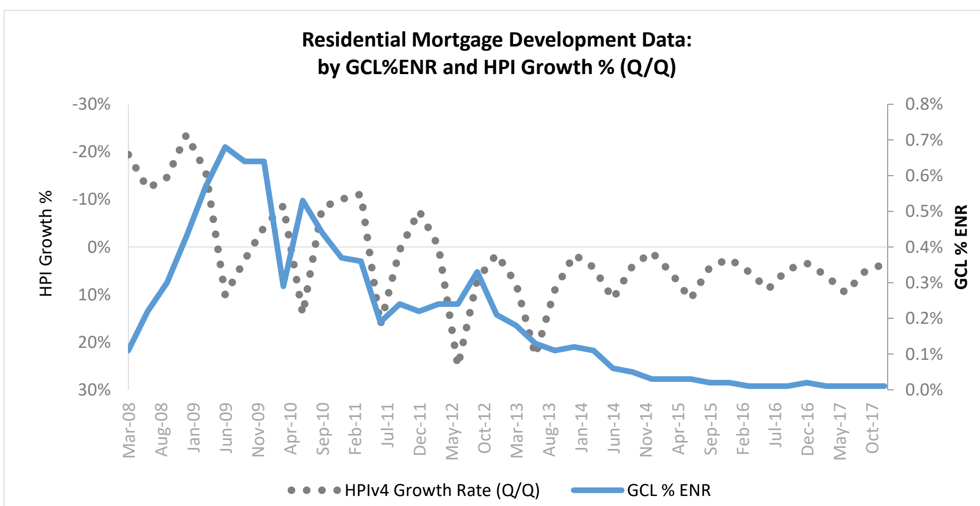
The portfolio dataset used in model development includes historical performance back to 2006 to 2017 covering both pre and post-financial crisis, covering the stress periods where portfolio experienced peak of the loss performance measured by GCL%ENR. In addition, it was during this stress period the overall US economy experienced the worst financial crisis since the great depression exhibiting consecutive quarters of negative GDP growth rate and observed the biggest drop in home prices measured in HPI index.

Figure 3.2.1.3 Residential Mortgage Historical Performance: GCL%ENR vs. Unemployment %



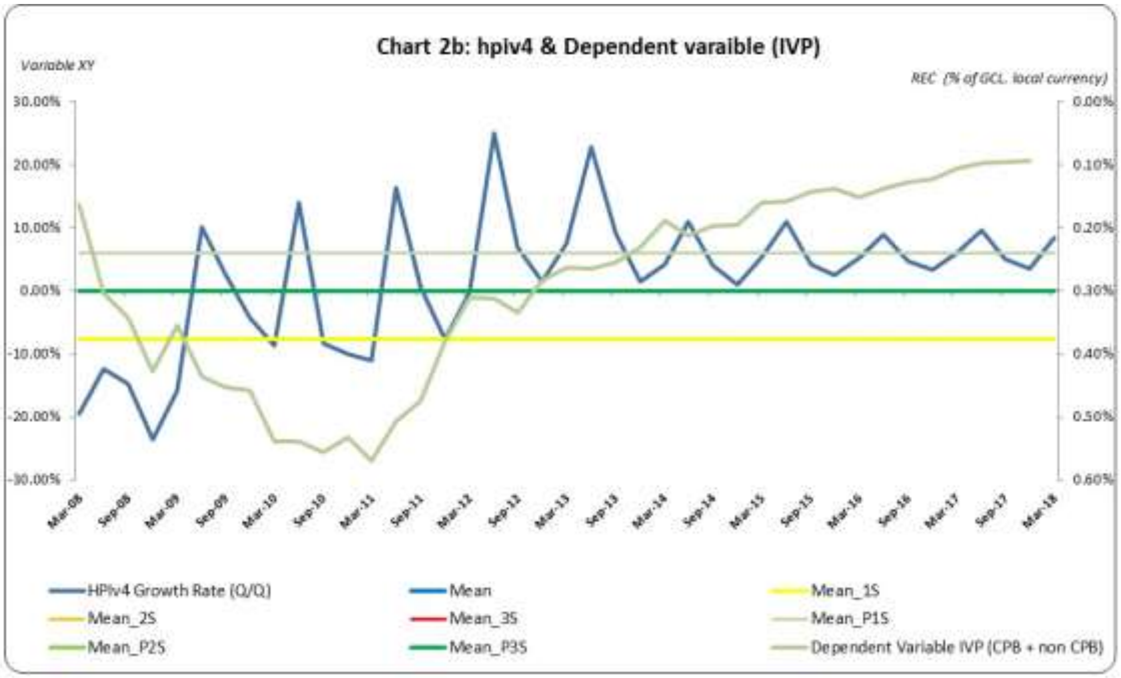
Stress period is define during 2008Q1 to 2010Q4, it was during this time period that unemployment rate (quarter over quarter change) started to rise in the beginning of 2009 period and hit its highest of 9.9%, during Q1 and Q2 of 2010. This coincides with the rising loss rate in the Residential Mortgage portfolio, reporting its peak value of 73 bps (basis points) starting at Q2 of 2009, and remains at high level until the Q4 2009.

Figure 3.2.1.4 Residential Mortgage Historical Performance: GCL%ENR vs. HPI Growth % (Q/Q)



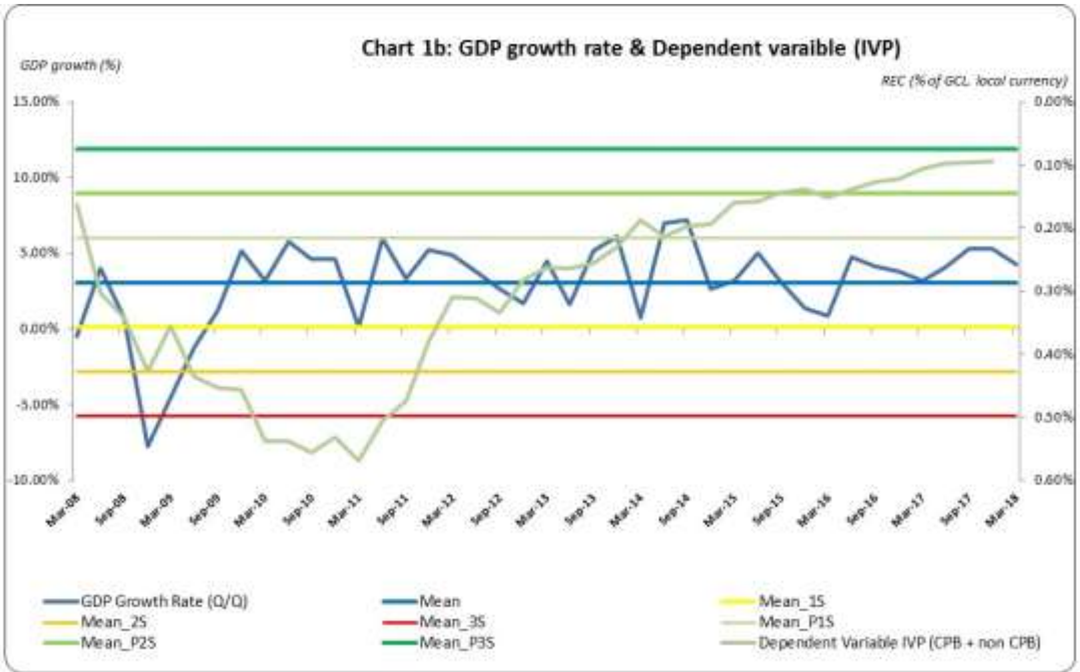
Moreover, similar stress was observed in terms of the drop in home prizes as measure by HPI (Home Prize Index). In Figure 3.2.1.4 portrays the relationship of the HPI growth rate (quarter over quarter growth) and Residential Mortgage Loss performance measure through GCL%ENR. It was during Q4 2008, when HPI registered the largest Q/Q drop of -23.6%, which followed by the sudden increase in losses in terms of GCL%ENR in the coming consecutive quarters in 2009.

Figure 3.2.1.5 Residential Historical Performance: Involuntary rate % vs. HPI Growth % (Q/Q)



The involuntary rate % (IVP) refers to the loss incidents mainly due to charge-off, foreclosure, short sales or loss pertaining to REO. In figure 3.2.1.5, IVP% is compared to the HPI quarter-over-quarter change showing the relationship between the IVP % and HPI growth %, during the stress period in 2009-2010. IVP rate started to drop beginning of 2009 during the same time when HPI growth rate drop the lowest.

Figure 3.2.1.6 Residential Mortgage Historical Performance: Involuntary rate % vs. GDP Growth % (Q/Q)



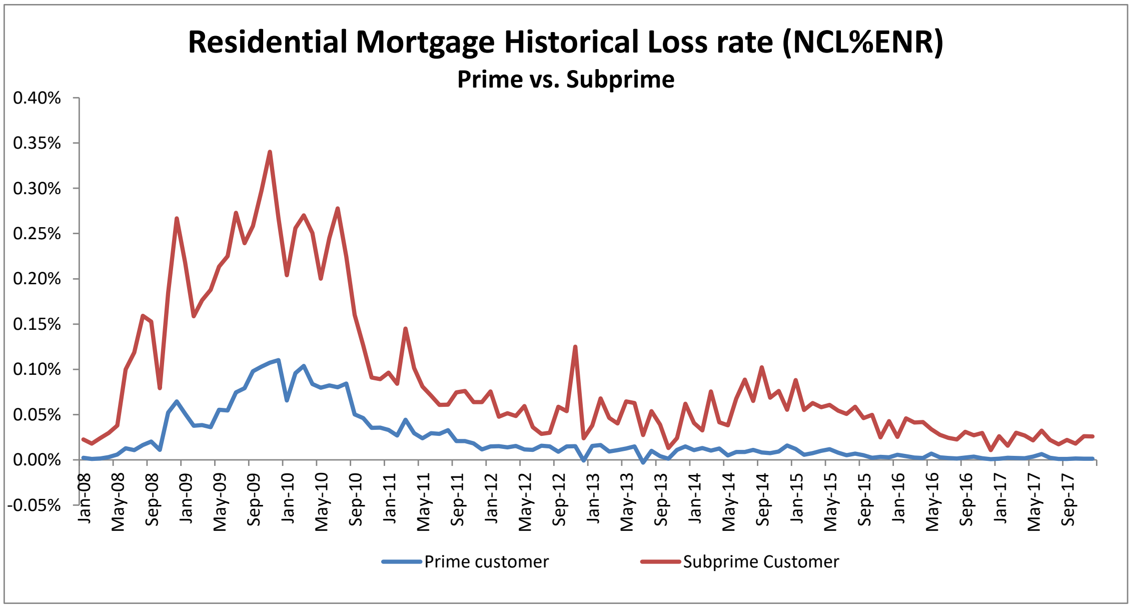
Comparing IVP % to GDP growth %, similar relationship has been observed, when GDP % were in the lowest rate in 2009. For further detail information, see attachment for 3.2.1 Gating Principle 3 Method A – BankMortgage (CMI) – Residential Mortgage.

Gating Principle 4: Review of the forecasted variable

- Behavior of the forecasted variable must be assessed with respect to patterns described in "MRM Library" as given in section 11.5 the Model Testing Guidance.

- If the expected portfolio pattern is missing from MRM Library or the developer disagrees with MRM's hypothesis, developer/Sponsor must provide their expectation of the portfolio behavior with sufficient rationale. Same should also be supported with data analysis. Model Sponsor is required to provide the information in the below table. Please refer to the Gating Principles as given in the Model Testing Guidance for more details.

Figure 3.2.1.4 Gating Principle Expected Behavior in Stress vs Non-Stress(Prime vs Sub-prime)



The portfolio behavior is consistent with the expected behavior define by MRM in the MRM Library. Subprime customers experienced higher losses (NCL%ENR) as the market goes through the stress period and during non-stress period, the tendency to default decreases, leading to lower loss. The Figure 3.2.1.4 above shows a comparison between prime and subprime customers through history during stress period of 2008-2010 and non-stress (recovery period) post 2010.

Table 3.2.1.5 Gating Principle Expected Behavior in Stress by Business Portfolio

RISK_SEGMENT	FILE_DT	CMI	CPB	Grand Total
PRIME	2008	0.01%	0.00%	0.01%
	2009	0.06%	0.00%	0.05%
	2010	0.06%	0.00%	0.05%
	2011	0.03%	0.01%	0.02%
	2012	0.02%	0.01%	0.02%
	2013	0.01%	0.01%	0.01%
	2014	0.01%	0.00%	0.01%
	2015	0.01%	0.00%	0.01%
	2016	0.00%	0.00%	0.00%
	2017	0.00%	0.00%	0.00%
PRIME Total		**0.02%**	**0.00%**	**0.02%**
SUBPRIME	2008	0.07%	0.00%	0.07%
	2009	0.18%	0.00%	0.18%
	2010	0.15%	0.01%	0.14%
	2011	0.06%	0.00%	0.06%
	2012	0.06%	0.01%	0.06%
	2013	0.04%	0.02%	0.04%
	2014	0.06%	0.02%	0.06%
	2015	0.06%	0.01%	0.06%
	2016	0.04%	0.01%	0.04%
	2017	0.03%	0.01%	0.03%
SUBPRIME Total		**0.09%**	**0.01%**	**0.09%**
Grand Total		**0.05%**	**0.00%**	**0.04%**

Similar to sub-prime vs prime, several other key risk drivers were tested during model development which pertain to CLTV,FICO,unemployment, HPI etc. All of these charts/tables are presented above as part of GP4 response or GP3 response. 95+%, resulted the highest loss incident during the stress period, peaking at 2009 Q4. Similarly, lowest FICO score band <620, exhibits the highest loss during the period of 2009Q4. All trends align with macro-economic trend and commonly perceived business intuition which states that borrower credit profile and correspondingly loss incidents(losses increase during stress) and macro-trends tend to deteriorate during stress and improve during non-stress. All of these attributes were added as separate modeled attributes or combined as interaction variables as part of the regression analyses. Further, a comparative assessment of portfolio composition change since the historical stress to recent times has been conducted within 6.4 of the MDD.

Table 3.2.1.4.1 Gating Principle Expected Behavior in Stress vs Non-Stress(FICO and CLTV)

NCL%ENR: Residential Mortgage By Business Portfolio				
FILE_DT	FICO Band	CMI	CPB	Grand Total
2008	1.<620	0.15%	0.00%	0.14%
	2.620 - 659	0.02%	0.00%	0.02%
	3.660 - 679	0.01%	0.00%	0.01%
	4.680 - 699	0.00%	0.00%	0.00%
	5.700 - 739	0.00%	0.00%	0.00%
	6.740+	0.00%	0.00%	0.00%
2008 Total		**0.04%**	**0.00%**	**0.04%**
2009	1.<620	0.32%	0.00%	0.31%
	2.620 - 659	0.14%	0.00%	0.13%
	3.660 - 679	0.05%	0.00%	0.04%
	4.680 - 699	0.02%	0.00%	0.01%
	5.700 - 739	0.01%	0.00%	0.01%
	6.740+	0.00%	0.00%	0.00%
2009 Total		**0.12%**	**0.00%**	**0.11%**
2010	1.<620	0.22%	0.03%	0.21%
	2.620 - 659	0.17%	0.03%	0.16%
	3.660 - 679	0.06%	0.00%	0.05%
	4.680 - 699	0.03%	0.00%	0.02%
	5.700 - 739	0.01%	0.00%	0.01%
	6.740+	0.01%	0.00%	0.01%
2010 Total		**0.11%**	**0.00%**	**0.10%**
2011	1.<620	0.10%	0.03%	0.10%
	2.620 - 659	0.07%	0.02%	0.07%
	3.660 - 679	0.03%	0.00%	0.03%
	4.680 - 699	0.02%	0.00%	0.01%
	5.700 - 739	0.01%	0.00%	0.01%
	6.740+	0.01%	0.00%	0.00%
2011 Total		**0.04%**	**0.01%**	**0.04%**
2012	1.<620	0.08%	0.03%	0.08%
	2.620 - 659	0.06%	0.02%	0.06%
	3.660 - 679	0.04%	0.02%	0.04%
	4.680 - 699	0.03%	0.01%	0.03%
	5.700 - 739	0.01%	0.00%	0.01%
	6.740+	0.00%	0.00%	0.00%
2012 Total		**0.04%**	**0.01%**	**0.03%**
2013	1.<620	0.06%	0.03%	0.05%
	2.620 - 659	0.05%	0.05%	0.05%
	3.660 - 679	0.01%	0.01%	0.01%
	4.680 - 699	0.02%	0.01%	0.02%
	5.700 - 739	0.01%	0.00%	0.01%
	6.740+	0.00%	0.00%	0.00%
2013 Total		**0.02%**	**0.01%**	**0.02%**
2014	1.<620	0.08%	0.02%	0.07%
	2.620 - 659	0.07%	0.08%	0.07%
	3.660 - 679	0.05%	0.00%	0.04%
	4.680 - 699	0.01%	0.03%	0.02%
	5.700 - 739	0.01%	0.00%	0.00%
	6.740+	0.00%	0.00%	0.00%
2014 Total		**0.03%**	**0.01%**	**0.03%**
2015	1.<620	0.08%	0.01%	0.07%
	2.620 - 659	0.07%	0.02%	0.07%
	3.660 - 679	0.05%	0.00%	0.04%
	4.680 - 699	0.02%	0.00%	0.02%
	5.700 - 739	0.00%	0.00%	0.00%
	6.740+	0.00%	0.00%	0.00%
2015 Total		**0.02%**	**0.00%**	**0.02%**
2016	1.<620	0.06%	0.02%	0.05%
	2.620 - 659	0.05%	0.01%	0.05%
	3.660 - 679	0.03%	0.01%	0.02%
	4.680 - 699	0.02%	0.00%	0.01%
	5.700 - 739	0.00%	0.00%	0.00%
	6.740+	0.00%	0.00%	0.00%
2016 Total		**0.01%**	**0.00%**	**0.01%**
2017	1.<620	0.04%	0.01%	0.03%
	2.620 - 659	0.05%	0.00%	0.04%
	3.660 - 679	0.01%	0.00%	0.01%
	4.680 - 699	0.01%	0.00%	0.01%
	5.700 - 739	0.00%	0.00%	0.00%
	6.740+	0.00%	0.00%	0.00%
2017 Total		**0.01%**	**0.00%**	**0.01%**
Grand Total		**0.05%**	**0.00%**	**0.04%**

NCL%ENR: Residential Mortgage By Business Portfolio				
FILE_DT	LTV Band	CMI	CPB	Grand Total
2008	1.< 50	0.04%	0.00%	0.03%
	2.50-74	0.06%	0.00%	0.05%
	3.75-79	0.03%	0.00%	0.03%
	4.80-89	0.03%	0.00%	0.03%
	5.90-95	0.03%	0.00%	0.03%
	6.>95	0.05%	0.00%	0.05%
2008 Total		**0.04%**	**0.00%**	**0.04%**
2009	1.< 50	0.34%	0.00%	0.26%
	2.50-74	0.21%	0.00%	0.18%
	3.75-79	0.09%	0.00%	0.08%
	4.80-89	0.07%	0.00%	0.06%
	5.90-95	0.06%	0.00%	0.05%
	6.>95	0.05%	0.00%	0.05%
2009 Total		**0.12%**	**0.00%**	**0.11%**
2010	1.< 50	0.34%	0.01%	0.26%
	2.50-74	0.22%	0.01%	0.19%
	3.75-79	0.08%	0.00%	0.07%
	4.80-89	0.06%	0.00%	0.05%
	5.90-95	0.04%	0.00%	0.04%
	6.>95	0.03%	0.00%	0.03%
2010 Total		**0.11%**	**0.00%**	**0.10%**
2011	1.< 50	0.11%	0.01%	0.09%
	2.50-74	0.09%	0.01%	0.08%
	3.75-79	0.03%	0.00%	0.03%
	4.80-89	0.03%	0.00%	0.03%
	5.90-95	0.02%	0.00%	0.02%
	6.>95	0.01%	0.00%	0.01%
2011 Total		**0.04%**	**0.01%**	**0.04%**
2012	1.< 50	0.06%	0.01%	0.05%
	2.50-74	0.06%	0.01%	0.05%
	3.75-79	0.03%	0.00%	0.02%
	4.80-89	0.05%	0.00%	0.04%
	5.90-95	0.03%	0.00%	0.03%
	6.>95	0.01%	0.00%	0.01%
2012 Total		**0.04%**	**0.01%**	**0.03%**
2013	1.< 50	0.01%	0.01%	0.01%
	2.50-74	0.02%	0.01%	0.02%
	3.75-79	0.02%	0.00%	0.02%
	4.80-89	0.03%	0.00%	0.02%
	5.90-95	0.03%	0.01%	0.02%
	6.>95	0.03%	0.01%	0.03%
2013 Total		**0.02%**	**0.01%**	**0.02%**
2014	1.< 50	0.01%	0.00%	0.01%
	2.50-74	0.03%	0.01%	0.03%
	3.75-79	0.04%	0.00%	0.03%
	4.80-89	0.04%	0.00%	0.04%
	5.90-95	0.04%	0.02%	0.04%
	6.>95	0.04%	0.04%	0.04%
2014 Total		**0.03%**	**0.01%**	**0.03%**
2015	1.< 50	0.01%	0.00%	0.00%
	2.50-74	0.02%	0.00%	0.01%
	3.75-79	0.03%	0.00%	0.02%
	4.80-89	0.05%	0.00%	0.04%
	5.90-95	0.05%	0.02%	0.05%
	6.>95	0.10%	0.01%	0.10%
2015 Total		**0.02%**	**0.00%**	**0.02%**
2016	1.< 50	0.00%	0.00%	0.00%
	2.50-74	0.01%	0.00%	0.01%
	3.75-79	0.02%	0.01%	0.01%
	4.80-89	0.04%	0.00%	0.03%
	5.90-95	0.04%	0.03%	0.04%
	6.>95	0.11%	0.08%	0.10%
2016 Total		**0.01%**	**0.00%**	**0.01%**
2017	1.< 50	0.00%	0.00%	0.00%
	2.50-74	0.00%	0.00%	0.00%
	3.75-79	0.01%	0.00%	0.01%
	4.80-89	0.03%	0.00%	0.02%
	5.90-95	0.03%	0.00%	0.03%
	6.>95	0.15%	0.00%	0.13%
2017 Total		**0.01%**	**0.00%**	**0.01%**
Grand Total		**0.05%**	**0.00%**	**0.04%**

Table 3: Review of forecasted variables

Portfolio behavior review	*Developer Response*	*Additional Comments*

1A - If Pattern is available in MRM Library does the Developer Agree with MRM Defined Behavior?	Yes	Developer agrees in the expected behavior in both stress and non-stress scenario that losses increase as market goes stress observing higher losses.
1B - If Pattern is NOT available in MRM Library define expected portfolio behavior in Stress and Non-Stress	N/A	
2 - Provide explanation to justify the Behavior in Stress and Non-Stress	Sub-prime segments customers are highly sensitive to market shock that directly affect the ability to pay. During stress period, the probability of default increases leading to higher loss. For non-stress environment, default, performance is more stable and lower losses are observed.	
3 – Does the data support the expected portfolio behavior as explained in 2?	Yes	Both macroeconomic variables GDP and HPI growth rate exhibit high correlation between the dependent variable (Cumulative loss in 6 moths rate), coinciding almost in the same period.

1 MRM Question – Model sponsor needs to provide supporting analysis that the loss trend is in line with that prescribed in MRM library on key segment such as CMI/CPB/SFO portfolio as well as on data of each model component (PD/LGD/DV logic/EAD)

Answer – Each model component is reviewed and examined carefully in the development period to ensure there is sufficient data for performance. As showed within the Gating Principles' [CAMU already provided the business/portfolio breakdown for RM (by business/FICO/LTV) in Gating Principles for the PD component of the model. Please refer to Gating Principles' for details] both RM PD model and LGD model satisfied the minimum sample size requirements in Gating Principle 1. The aggregated GCL in terms of unit and of amount in Table 3.2.1.4 provide further support that the data follows the economic cycle and is sufficient for modeling. The DV Logic data captures the log ratio of end appraisal value and beginning appraisal value at stress and non-stress periods. As illustrated in Table 3.2.1.5 and Figure 3.2.1.2, the data is also sufficient for modeling and follows the economic cycle.

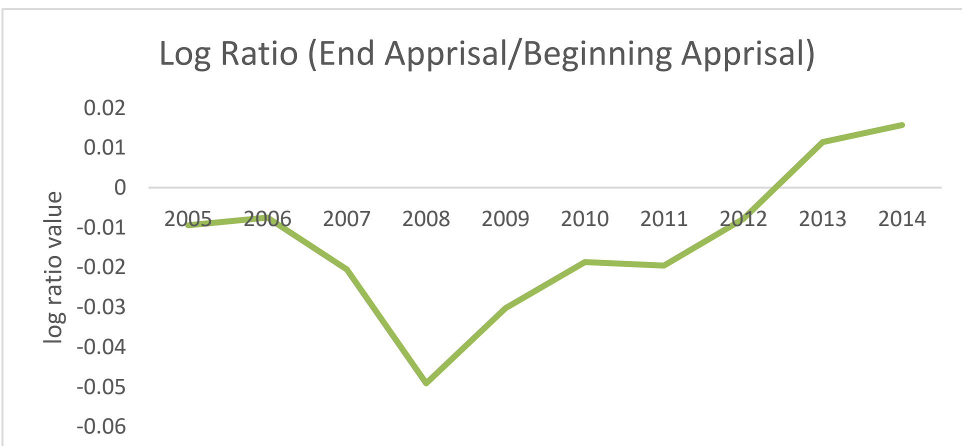
Table 3.2.1.3 Loss in LGD Model

yr	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Total Loss Unit	46,912	71,768	59,719	46,192	31,714	18,935	11,573	7,815	5,165	2,204
Total Loss Amount	3,248,450,999	5,358,111,961	4,202,218,976	3,242,233,454	2,108,485,234	1,106,032,547	611,629,454	401,894,423	262,711,973	107,659,291
Principal Balance at Year End	$590,002,933	$581,613,580	$431,976,162	$324,607,427	$179,237,895	$93,433,646	$75,087,522	$41,650,502	$23,619,706	$32,497,325

Table 3.2.1.4 Appraisal Values in DVM Model

Yr	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Total Number of Observations	20,696	34,111	43,587	64,676	127,565	186,537	191,450	180,187	147,047	53,379
early appraisal value	102,205	117,782	154,076	183,948	185,752	189,914	186,725	186,415	184,056	198,683
appraisal value at year	100,002	115,766	146,955	164,278	173,258	181,903	178,512	183,079	188,961	206,005
dependent variable: ratio in log	-0.0095	-0.0075	-0.0206	-0.0491	-0.0302	-0.0187	-0.0195	-0.0078	0.0114	0.0157

Figure 3.2.1.2 Log Ratio of Appraisal values



Gating Principle 5: Portfolio Segmentation

- To show that the segmentation approach has appropriate granularity (i.e., segments have sufficient data to develop robust and testable estimates that capture different underlying portfolio characteristics) given the modeling objective of forecasting losses/balances under normal and stressed macroeconomic environments, please provide:
  - Details on segmentation schemes explored and how the chosen segmentation scheme meets the modeling objective.
  - Business rationale for the segmentation, which could either be business requirements driving a non-statistical segmentation or the business intuition for the segmentation variables of a statistical segmentation
    - Evidence of what leads you to believe that developing a model based on the chosen n this segmentation approach is feasible (e.g., number of defaults per segment)
    - For statistical segmentations, discussion of the trade-offs compared to alternative segmentations such as an MA with appropriate explanatory variables but no segmentation
    - For statistical segmentation, if available, analysis to justify differences between segments (e.g., central tendency and / or dispersion of distributions, risk factors, sensitivities to common risk factors)

  *[*Required: Please provide response here. Even if there is no segmentation, please provide the rationale for no segmentation.]*

  The segmentation scheme used in the Gating Principle 5 template is based on non-statistical approach and mainly aligns with Business Sponsor CCAR Reporting. The segmentation used specifically for the Gating Principle template is high-level portfolio segments (by business, and product) and delinquency buckets (Current, 30-60 bucket and 90+ buckets), which are the commonly used metrics by the business sponsor. The data provided by the Risk MIS Reporting team, who is responsible for all external reporting including CCAR. This makes sure that the data went through the same toll gating process and data has been verified to be consistent with the other external reports. The original segments used in the gating principle are not exactly the same, compared to the actual model segmentation using more detail statistical approach. The main rationale behind the chosen segmentation is to provide a benchmark and guidance for further exploratory segmentation analysis, to make sure business reports are consistent with other reporting tools and for model sponsors can provide high-level strategic business plan in the major key segments and risk indicators as reference point for more detail segmentation. Detail segmentation using statistical approach is discussed in detail in section 5.1.3 which is mainly based on decision tree approach and other statistical testing approaches.

- Please articulate which variables influence the dependent variable (these could include macroeconomic variables, behavioral variables, etc.), and justify using any of the options below:
    - Regulatory guidance and industry literature
    - Qualitative description of business justification
    - Statistical evidence, if available from current or prior development efforts

*[*Required: Please provide response here]*

For details on variables influencing the dependent variable, please refer to chapter 5 Model Specification.

- If there were other input challenges, please describe
  - The CPB portfolio is slightly different credit profile which is more targeted to affluent customers and with deeper client relationship with other Bank products or assets. Overall performance is better compared to CMI residential mortgage, but have potentially significant due to the exposure of portfolio.
  - The Service for Others (SFO) was not part of the model development data, however used in Method A models as borrow model.

*[*Optional: Please provide response here]*

Appropriateness of Coping:

- Acknowledging this limitation of the CPB, the coping mechanism used is combine and blend their credit performance together in terms of modeling the overall performance of the residential mortgage.
- The SFO data were never considered in the modeling development, however, the SFO portfolio was validated to perform similar behavior during the backtesting results. This one of the coping mechanism to justify that the model is appropriate for the usage in the SFO portfolio. Additionally, since the recent development and business strategy of exiting the servicing business as result of the Mortgage Transformation, it is expected that portfolio will soon to further liquidate and reduce further exposure which provide a practical approach and business rational for not building a separate model for SFO loans.

[Please provide the documentation to demonstrate the appropriateness of Coping. All the details on appropriateness of coping strategy must be provided as required by the Gating Principle guidance].

Technique:

- If there were any challenges with technique in current or prior development efforts, please describe (e.g., variables were not predictive or had a counterintuitive relationship with model output)

*[*Optional for Pre-MEA: Please provide response here]*

Output:

- If there were any output challenges in current or prior development efforts, please describe

(e.g., statistical tests exceeded prescribed error thresholds, or outputs under stress events do not match business hypotheses)

*[*Optional for Pre-MEA: Please provide response here]*

Other Challenges:

- If there were any other challenges, please describe

  *[*Optional: Please provide response here]*

If there were challenges above, please provide evidence that appropriate mitigating actions were attempted in order to overcome the challenges and explain whether these actions were successful.

Below are some examples of mitigating actions:

- For challenges related to inputs: Supplementing existing, internal datasets with external data, using a proxy portfolio, building a base model with a stress overlay
- For challenges related to technique: Applying different type of regression (e.g., fractional logic rather than linear)
- For challenges related to output: Applying caps and/or floors to ensure outputs remain within expected ranges

- As mentioned in Section 1.1 of the MDD, the RM model suite has been leveraged for the SFO portfolio too. Appropriate justification for borrowing this model has been described in Section 1.1 of the MDD. Further justification for the combining of the CMI and CPB portfolios have been analyzed both statistically and rationally (decision tree engine and commonly perceived business intuition). Please refer to Section 5.1.3 of the MDD for additional details. Additional evidence capturing the differentiated performances for CMI vs CPB across stress and non-stress periods have been detailed in the Variable Selection Process in Section 5.1.4. For the DV Logic related coping strategy across recent and stress periods, please refer to Section 5.1.2 of the MDD.

Coping Strategy – CMI vs CPB - As mentioned in Section 1.1 of the MDD, both-business entities (CMI, CPB) have demonstrated very specific idiosyncratic features in terms of loan quality and sensitivity to upheavals in the broader economy. While the CMI portfolio always had a rich history of data across stress and non-stress, the CPB portfolio did not experience a typical stress period per se, as the CPB clients have always demonstrated muted sensitivity to macro-economic factors. To provide some context, in particular, there were only 20 residential mortgage loans with $9MM lifetime losses during the recession 2008-2009. Obviously, such low volume of defaults/losses implied that a) Gating Principles that establish the sufficiency and suitability of data for the built of statistical models could not be met; b) there could be no statistical model that can provide robust predictions for such low volume of losses. As such, the following steps were taken as appropriate coping strategies for the effective modeling of the CPB portfolios within the realm of the Method A RM model suite-

PD Model – The CPB portfolio was combined with the CMI portfolio since not much differentiated sensitivity were observed between these two specific portfolios based on the decision tree analysis. Please refer to Section 5.1.3 for additional details on the decision-tree. However, specific balance related attributes were introduced for the CMI portfolio to track the CMI specific balance effects in the model and differentiate CMI loans performance from CPB loans. Please refer to the model specification listing in Section 5.1.4 for additional details.

LGD model – Since the CPB loans had extremely low volume of losses, which also didn't align with the macro-economic intuition (stress losses were lower than recent period), it was considered prudent to utilize a lookup table approach for the CPB loans. The strategy deployed slightly differed between the base and stress scenarios. For the first lien loans - base scenario, an empirical lookup table was created leveraging the historical loan performance data from 201101 to 201706 to forecast the base losses. In order to model the stress scenario, a twofold approach was utilized. First, as a coping mechanism, the CMI first lien portfolio was chosen to serve as a proxy for estimating CPB first lien stress losses. The CMI portfolio was primarily chosen since it had a rich history of data with sufficient sensitivity to business cycles. The CMI empirical loss curve was plotted to gauge the level of peak losses. The empirical data revealed that there is usually a time lag in the loss recognition due to the delays in the foreclosure/bankruptcy processes. So while the historical stress period persisted from 2008-2009, the actual losses peaked from 2009-2011. This peak point of the loss curve was then utilized to calculate a Stress Multiplier. The multiplier was calculated by taking the ratio of the peak period losses to the recent period losses within the CMI portfolio. This multiplier was subsequently used to calculate the add-on losses over and above the base losses for the CPB portfolio.

For all model forecasting and backtesting needs, the derived LGD multiplier was utilized for all stress scenarios. For the baseline scenarios, the empirical lookup table was utilized to estimate the expected losses. Since there were only 50 accounts under second liens that experienced losses for the CPB portfolio, it was considered prudent to just assign full losses (loss rate=1) to these second liens for all forecasting and backtesting needs, including both stress and base scenarios. Please refer to Section 5.1.2 of the MDD for additional details on the lookup logic. Additional evidence capturing the differentiated performances for CMI vs USRB vs CPB across stress and non-stress periods have been detailed in the Variable Selection Process in Section 5.1.4.

DV Logic Related Coping Strategy - The Method A model suite utilized a newly developed DV lookup logic that replaces the prior year's DV model. The new haircut logic is based on data that spans across both stress and non-stress periods and is aligned with the property valuation methodologies as utilized by Bank's Collateral Management team. Please note while the DM data is available starting 2010, the stress period utilized the DRI data. The haircut ratios were tested across time (stress + non-stress) for both PDV [prior distressed valuation available] and NPDV [no prior distressed valuation available] segments across a multitude of risk drivers (loan level) that potentially impact property values. Please refer to attachments –'5.1.2 pivot_PDV_Haircut_newData_HPIV4byQtr.xlsm' and '5.1.2 pivot_NPDV_Haircut_newData_HPIV4byQtr.xlsm' for details on the tested dimensions. Additionally, as part of the coping strategy (justifying the suitability of using both the DM and DRI data sources for the calculation of the Haircut Logic), the attachment –'5.1.2 DRI_vs_DM_On_PDV.xlsb' presents the side-by-side comparison of the haircut ratio as derived on DRI and DM. It shows that the haircut ratio based on DM data is close to DRI data when both datasets are available. So although the DM data is not available for the stress period, it was considered prudent to borrow the haircut ratio as obtained from the DRI data.

As mentioned in Section 1.1 of the MDD, the RM model suite has been leveraged for the SFO portfolio too. Appropriate justification for borrowing this model has been described in Section 1.1 of the MDD.

3.2.2 Describe the progression of MAs attempted, and highlight why an approach was rejected. Then describe how you chose to refine the MA. Highlight the challenges faced, (e.g., a direct model of the reference portfolio could not be estimated because the number of necessary observations were insufficient) and provide evidence of appropriate mitigating actions to overcome the challenges before proceeding to another alternative Modeling Approach.

3.2.2a Alternative for the PD Model - Non modeled transitions

As described in Section 3.1, CAMU took a systematic way to estimate the non-modeled transitions. During the model development period, leading up to the finalization of the preferred approach for modeling the non-modeled transitions, several alternative approached were tested which ultimately were dropped due to concerns around uplift of model performance. Some of the strategies that were tried are listed as below-

Alternative Approach # 1

Combine neighboring cells to mitigate volume limitation. One of the approaches that was tested during the initial planning stages was to combine transitions which exhibited similar sensitivities to key risk drivers & to macroeconomic variables with the addition of appropriate dummy effects to capture the nuances of each transition. This was considered a viable alternative approach and had been used by the Method B model suite (2017 Challenger model). However when it came around to evaluate this effectivity of this approach using model's backtesting, it was noted that the backtesting itself required the re-distribution of the collated loans which again became a modeling challenge, given the already limited volume associated with these transitions. Incorrect loan allocation during backtesting could affect the marginal predictions of the loans in the specific buckets which could potentially impact model's performance. The same concern was also raised by the model reviewers' on the Method B model suite and CAMU used this experience as lesson learnt and considered prudent to drop this approach from the scope.

Alternative Approach # 2

The second alternative to the preferred approach was to leverage a constant rate for estimating the non-modeled (rare) transitions, across stress and non-stress periods. This approach had been used by the Method A model suite in prior year's models. However, using one averaged constant roll rate over the entire time horizon couldn't fully differentiate the extreme differences in actual performances between stress vs recent periods, thus potentially affecting the model's fit in recent time –periods which has witnessed periods of steadily rising HPI with declining unemployment. Again based on the deep-dive analyses conducted this year on the rare transitions, it was noted that more often than not these rare transitions are very volatile over time with significant changes from stress to non-stress. This was actually observed in prior year's model results and was cited as a limitation. Given the evidence from past submission and subjective recommendation by model reviewers', there was a deliberate attempt made to steer away from this approach in this year's process.

Alternative Approach #3

A third approach that was tried and led to the creation of a look-up table for specific shorted timeframes such that the model's fit could be improved for these timeframes. However, it is to be noted that all econometric models are built to capture a broad horizon of economic movement. Deviations used to address specific period's trends not only breaks the modeling data thinner but can also cause potential over-fitting leading to the fallacy of over-causation.

Given the flaws associated with the alternative approaches, the preferred approach was considered the best approach that would help address the prior model's limitation in a systematic, tested way.

3.2.2b DV Logic

As mentioned in the preceding section, the main alternative for the new DV Logic is the existing/prior DVM model. The new DV logic has been conceptualized based on feedback on the prior modeling approach and CAMU's commitment to continuously develop and adopt new processes and functionalities based on the evolving trends within the market. Having established the premises of the new logic, CAMU has provided the following in-depth analysis that demonstrates and justifies the selection of this proposed logic for the 2019 CCAR process.

Provided below is the detailed comparison between the proposed DV logic and the existing DVM, via mathematic derivation and excel examples (Please refer to attachment 3.2.2 DV Sensitivity Comparisons.xlsx). The two methods will be referred to as 2019 DV Logic and 2017 DVM respectively in the following discussion.

Discussion below derives the distressed property value, it's month over month change for both the NPDV and PDV segments on both methods. It shows that the 2019 method leads to same macroeconomic sensitivity as the 2017 DVM on the NPDV segment, and has a higher sensitivity on the PDV segment.

Existing DVM

3.2.2b.1 NPDV Segment

On the NPDV segment, the prior model estimates the DV of a certain month as:

$$\mathrm{DV(t)} = \exp\left(\mathrm{a_{npdv}} + \mathrm{b_{npdv}} * \ln\frac{\mathrm{HPI(t)}}{\mathrm{HPI(Orig)}} + \mathrm{c_{npdv}}^{\mathrm{MOB}}\right) * \mathrm{Orig\_Prop\_Amt}$$

.... 1)

where $\mathrm{a_{npdv}} = -0.192630, \mathrm{b_{npdv}} = 1, \mathrm{c_{npdv}} = -0.0012237.$

$\mathrm{c_{npdv}}$ is so small that $\mathrm{c_{npdv}^{MOB}} \to 0$, hence 1) can be written as:

$$\mathrm{DV(t)} = \exp(\mathrm{a_{npdv}}) * \exp(\mathrm{c_{npdv}^{MOB}}) * \exp\left(\ln\left(\frac{\mathrm{HPI(t)}}{\mathrm{HPI(Orig)}}\right)\right) * \mathrm{Orig\_Prop\_Amt}$$

$$\approx \exp(\mathrm{a_{npdv}}) * \frac{\mathrm{HPI(t)}}{\mathrm{HPI(Orig)}} * \mathrm{Orig\_Prop\_Amt}$$

$$= \exp(-0.192630) * \frac{\mathrm{HPI(t)}}{\mathrm{HPI(Orig)}} * \mathrm{Orig\_Prop\_Amt}$$

$$\approx 0.8248 * \frac{HPI(t)}{HPI(Orig)} * Orig\_Prop\_Amt$$

.... 2)

PCO logic also uses the month-over-month DV change to determine the incremental partial charge-offs, which can be written as below:

$$DV(t) - DV(t-1) = 0.8248 * Orig\_Prop\_Amt * \frac{HPI(t) - HPI(t-1)}{HPI(Orig)}$$

.... 3)

Its sensitivity by comparing the DV value based on different HPI scenarios at the same month *t* can be expressed as:

$$\frac{DV(stress,t)}{DV(base,t)} = \frac{0.8248 * \frac{HPI(stress,t)}{HPI(Orig)} * Orig\_Prop\_Amt}{0.8248 * \frac{HPI(base,t)}{HPI(Orig)} * Orig\_Prop\_Amt}$$

$$= \frac{HPI(stress,t)}{HPI(base,t)}$$

.... 4)

And sensitivity of month-over-month DV changes based on different HPI scenarios is:

$$\frac{DV(stress,t) - DV(stress,t-1)}{DV(base,t) - DV(base,t-1)}$$

$$= \frac{0.8248 * Orig\_Prop\_Amt * \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Orig)}}{0.8248 * Orig\_Prop\_Amt * \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Orig)}}$$

$$= \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Base,t) - HPI(Base,t-1)}$$

.... 5)

3.2.2b.2 PDV Segment

Formula of DV in the PDV segment in the prior DVM is:

$$DV(t) = \exp\left(a_{pdv} + b_{pdv} * ln\frac{HPI(t)}{HPI(time\ of\ last\ DV)} + c_{pdv}{}^{month\ since\ last\ DV}\right) * Amt\ of\ last\ DV$$

.... 6)

where $a_{pdv} = -0.037598, b_{pdv} = 1, c_{pdv} = -0.0010924.$

c_{npdv} is so small that $c_{npdv}^{month\ since\ last\ DV} \to 0$, hence 6) becomes:

$$DV(t) = \exp(a_{pdv}) * \exp(c_{pdv}^{month\ since\ last\ DV}) * \exp\left(\ln\left(\frac{HPI(t)}{HPI(time\ of\ last\ DV)}\right)\right) * Amt\ of\ last\ DV$$

$$\approx \exp(a_{pdv}) * \frac{HPI(t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV$$

$$= \exp(-0.037598) * \frac{HPI(t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV$$

$$\approx 0.9631 * \frac{HPI(t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV$$

.... 7)

Month-over-month DV change to determine the incremental partial charge-offs can be written as:

$$DV(t) - DV(t-1) = 0.9631 * Amt\ of\ last\ DV * \frac{HPI(t) - HPI(t-1)}{HPI(Orig)}$$

.... 8)

Its sensitivity by comparing the DV value based on different HPI scenarios at the same month t can be expressed as:

$$\frac{DV(stress,t)}{DV(base,t)} = \frac{0.9631 * \frac{HPI(stress,t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV}{0.9631 * \frac{HPI(base,t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV}$$

$$= \frac{HPI(stress,t)}{HPI(base,t)}$$

.... 9)

The sensitivity of month-over-month DV changes based on different HPI scenarios is:

$$\frac{DV(stress,t) - DV(stress,t-1)}{DV(base,t) - DV(base,t-1)} = \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Base,t) - HPI(Base,t-1)}$$

.... 10)

2019 DV Logic

3.2.2b.3 NPDV Segment

In the new method, DV of a certain month is:

$$DV(t) = \frac{HPI(t)}{HPI(Orig)} * Orig\_Prop\_Amt * (1 - haircut_{npdv})$$

.... 11)

where $haircut_{npdv}$ depends only on loan level static features such as occupancy type or lien position.

Month-over-month DV change to determine the incremental partial charge-offs then can be written as:

$$DV(t) - DV(t-1) = (1 - haircut_{npdv}) * Orig\_Prop\_Amt * \frac{HPI(t) - HPI(t-1)}{HPI(Orig)}$$

.... 12)

So sensitivity by comparing the DV value based on different HPI scenarios at the same month is:

$$\frac{DV(stress,t)}{DV(base,t)} = \frac{\frac{HPI(stress,t)}{HPI(Orig)} * Orig\_Prop\_Amt * (1 - haircut_{npdv})}{\frac{HPI(base,t)}{HPI(Orig)} * Orig\_Prop\_Amt * (1 - haircut_{npdv})} = \frac{HPI(stress,t)}{HPI(base,t)}$$

.... 13)

The sensitivity of month-over-month DV changes based on different HPI scenarios is also:

$$\frac{DV(stress,t) - DV(stress,t-1)}{DV(base,t) - DV(base,t-1)}$$
$$= \frac{(1 - haircut_{npdv}) * Orig\_Prop\_Amt * \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Orig)}}{(1 - haircut_{npdv}) * Orig\_Prop\_Amt * \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Orig)}}$$
$$= \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Base,t) - HPI(Base,t-1)}$$

………………………… 14)

3.2.2b.4 PDV Segment

On the PDV segment, distressed property value is

$$DV(t) = \frac{HPI(t)}{HPI(,time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1 - haircut_{pdv})$$

………………………… 15)

where $haircut_{pdv}$ = 0.1 for stress scenario and 0 for base scenario.

Month-over-month DV change to determine the incremental partial charge-offs is:

$$DV(t) - DV(t-1) = \frac{HPI(t) - HPI(t-1)}{HPI(,time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1 - haircut_{pdv})$$

………………………… 16)

So sensitivity by comparing the DV value based on different HPI scenarios at the same month *t* is:

$$\frac{DV(stress,t)}{DV(base,t)} = \frac{\frac{HPI(stress,t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1-0.1)}{\frac{HPI(base,t)}{HPI(time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1-0)}$$

$$= \frac{HPI(stress,t)}{HPI(base,t)} * 0.9$$

.... 17)

The sensitivity of month-over-month DV changes based on different HPI scenarios is:

$$\frac{DV(stress,t) - DV(stress,t-1)}{DV(base,t) - DV(base,t-1)}$$

$$= \frac{\frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(,time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1-haircut_{pdv})}{\frac{HPI(base,t) - HPI(base,t-1)}{HPI(,time\ of\ last\ DV)} * Amt\ of\ last\ DV * (1-haircut_{pdv})}$$

$$= \frac{HPI(stress,t) - HPI(stress,t-1)}{HPI(Base,t) - HPI(Base,t-1)} * 0.9$$

.... 18)

Based on the above mathematic derivation, the following conclusion can be made on the NPDV segment:

a. In any given month, based on the same HPI scenario, ***the DV predicted by the 2019 DV Logic is always*** $(\mathbf{1} - \boldsymbol{haircut_{npdv}})/\mathbf{0.8248}$ ***times the DV predicted by 2017 DVM model in any given month***, according to equations # 2) and 11).
The table below highlights the ratio between DV based on 2019 DV Logic and existing DV model by taking into account the value of $haircut_{npdv}$ as proposed in previous section. When considering only the owner-occupied non-first lien, using the new method results in a 5.5% higher estimation for the property values compared to the prior DVM model. In the other three segments, the DV based on new method is always more conservative.

Table 3.2.2.1 Ratio between DV based on 2019 DV Logic and 2017 DVM on NDPV segment

Ratio	first lien	Non-first lien

Owner-Occupied	0.970	1.055
Not Owner-Occupied	0.885	0.946

b. As shown in equations 3) and 12), ***in both the DV methods, the month-over-month DV change is always proportional to the month-over-month HPI appreciation*** $\boldsymbol{HPI(t) - HPI(t-1)}$***.*** If HPI increases by 1% in 1-month, DV also increased by 1%. By comparing 3) and 12), it can also conclude that the ratio of absolute change of month-over-month DV value based on new method and that based on existing DV model is also $(1 - \mathrm{haircut_{npdv}})/0.8428$.

c. By comparing expression 4) and 13) it can be concluded that ***the two methods demonstrate similar sensitivity to changes in HPI scenarios***. For example, in any given month, if HPI in the stress scenario is 5% lower than the HPI in the base scenario, then the DV predicted for both the prior model and the new logic is 5% lower for the stress scenario compared to base.

d. 5) and 14) indicates that ***sensitivity of month-over-month DV change to different monthly HPI change scenarios is also the same in both the prior model and the new logic.***

Similarly, for the PDV segment, we can conclude that:

e. In any given month, based on the same HPI scenario, the ***DV predicted by the 2019 DV Logic is always*** $\boldsymbol{(1 - haircut_{pdv})/0.9631}$ ***times the DV predicted by the existing DVM model in any given month***, according to 7) and 15).
Note that for the PDV segment, 2019 DV Logic proposed different haircut ratios for different scenarios. $haircut_{pdv}$ equals 0 and 0.1 in the base and stress cases respectively.
Table below gives the ratio between DV based on 2019 DVM and 2017 DVM by taking into the above haircut value.
As shown in the table below, on the PDV segment, the newly proposed DV formula is more conservative than the existing DV model, as its predicted property value is only 93% of the value based on the existing model. On the other hand, the newly proposed DV method gives higher DV estimation than the existing model under the base scenario.

Table 3.2.2.2 Ratio between DV based on 2019 DVM and 2017 DVM on DPV segment

	Stress	**Base**
Ratio	0.934	1.038

f. 8) and 12) also indicate that ***the month-over-month DV change is always proportional to HPI appreciation*** $\boldsymbol{HPI(t) - HPI(t-1)}$ ***in both methods.*** Ratio of absolute change of month-over-month DV value based on 2019 DVM and that based on existing DV model is also $(1 - haircut_{pdv})/0.9631$.

g. According to 9) and 17), as well as 10) and 18), ***sensitivities;*** by comparing the absolute DV value in a given month, or by comparing the month-over-month DV value change under different HPI scenario, ***is always about (1/0.9-1)≈11% higher in 2019 DVM than in the existing DV model.***

In the attached excel, we made up a pseudo loan which was originated in Jan. 2012. Its property value at origination was 200K. Assuming this is an owner-occupied property not having any prior distressed property valuation and the loan associated with it is a first lien. Starting from Jun 2017, we calculated its monthly DV based on the proposed DV method and the existing DVM, under base and stress HPI scenarios.

For this NPDV loan, in the worksheet we compared –

- DV on the same month, using same HPI scenario, but based on different DV methods, which shows the value based on the new method is always 97% of that based on the existing DVM, which is consistent with table 2 above.
- Sensitivity to change of HPI scenario, based on different DV method, which shows the two models responses to the change of HPI in the same direction and with the same magnitude.

Another example was created for a pseudo loan in the PDV segment. Its prior distressed property value was 200K in Jan. 2012. The attached worksheet in tab "PDV" shows that –

- Under base scenario, the DV forecasted based on the newly proposed method is about 4% higher than that based on the existing DV model, whereas under the stress scenario, DV is about 6% lower. The ratios are consistent with numbers in table 3.
- Accordingly, the newly proposed DV method increases the DV model HPI sensitivity by about 11%.

The detailed example illustrating the above example can be found in attachment titled '3.2.2b DV Sensitivity Comparisons.xlsx.' Presented below are the comparisons of the 2018 DVM model with the 2019 DV Logic for the medium and long-term horizons from the sensitivity runs using 201806 snap date, respectively. For additional details on the sensitivity results, please refer to attachment – '3.2.2b DV_sensitivity_comparison.xlsx'. As expected and discussed earlier, using the new DV logic, uplifts the model's sensitivity by ~10%.

Table 3.2.2.3 Sensitivity Comparison -2018 DVM model vs 2019 DV Logic(27 month)

Scenario	2018 Model		2019 Model	
	Average DV ($)	% Change	Average DV ($)	% Change

Base (c01)	698,181	N/A	737,860	N/A
Stress (c08)	512,230	26.6%	525,212	28.8%

Table 3.2.2.4 Sensitivity Comparison -2018 DVM model vs 2019 DV Logic (60 month)

Scenario	2018 Model		2019 Model	
	Average DV ($)	% Change	Average DV ($)	% Change
Base (c01)	759,253	N/A	802,3453	N/A
Stress (c08)	520,277	31.5%	527,475	34.3%

3.2.2c Alternative for the LGD Model

As described in the Section 3.1 of the MDD, the LGD framework for the first lien loans has been redeveloped from modeling losses based on their disposition types to modeling losses based on their outcome (full vs partial vs zero). The new framework would leverage a two-stage approach towards modeling losses wherein the first stage segments the losses based on their outcome [full, partial and zero losses] and the second stage models the loss severity for the partial losses only. The main alternative to this new framework is the prior modeling approach for the first lien loans. As mentioned in the earlier section, the prior modeling approach was associated with many fallacies which led to its replacement with this new approach. This new framework not only alleviates the shortcomings of the prior framework but also aligns the modeling approaches across loan liens.

3.2.2.1 Alternative Modeling Approach #1

*[*Required only if the preferred Modeling Approach identified in Section 3.1 is not the final approach chosen and documented:*

Please describe Alternative Modeling Approach #1.

Please explain why Alternative Modeling Approach #1 was the next best option after the preferred Modeling Approach

If Alternative Modeling Approach #1 was not feasible please explain the challenges related to input (if different from what is outlined above), technique and output. This explanation of challenges is optional for pre-MEA.

Provide evidence that appropriate mitigating actions were attempted to overcome the challenges before proceeding to Modeling Approach #2] , or alternatively, if the development of *the preferred Modeling Approach has been deprioritized, please note that accordingly.*

*[*Required for Pre-MEA: Please provide response here; if the preferred approach was not deprioritized please say N/A]*

The modeling approaches discussed in Section 3.1 for the Method A RM PD and LGD model constitute the final approach and have been proven to be the best available for Method A suite of models, based on the alternatives discussed in MDD Section 3.2.2.

3.2.2.2 Alternative Modeling Approach #2

*[*Required only if Alternative Modeling Approach #1 in Section 3.2.2.1 is not the final approach chosen and documented:*

Please describe Alternative Modeling Approach #2.

Please explain why Alternative Modeling Approach #2 was the next best option after Alternative Modeling Approach #1

If Alternative Modeling Approach #2 was not feasible please explain the challenges related to input (if different from what is outlined above), technique and output. This explanation of challenges is optional for pre-MEA.

Provide evidence that appropriate mitigating actions were attempted to overcome the challenges before proceeding to Modeling Approach #3] , or alternatively, if the development of the preferred Modeling Approach has been deprioritized, please note that accordingly.

*[*Required for Pre-MEA: Please provide response here; if the preferred approach was not deprioritized please say N/A]*

The modeling approaches discussed in Section 3.1 for the Method A RM PD and LGD model constitute the final approach and have been proven to be the best available for Method A suite of models, based on the alternatives discussed in MDD Section 3.2.2.

3.2.2.3 Alternative Modeling Approach #[]

The modeling approaches discussed in Section 3.1 for the Method A RM PD and LGD model constitute the final approach and have been proven to be the best available for Method A suite of models, based on the alternatives discussed in MDD Section 3.2.2.

4. Model Data

[Documentation should include discussion of historical data set construction, including data sources, adjustments to the data set, and documentation validating the use of any external data.

The purpose of this chapter is to provide an understanding of the data used for modeling, while also providing evidence that data weaknesses and limitations have been thoroughly considered and can be justified. If data was not used as part of the Modeling Approach, please write N/A and explain why data was not used or questions are not applicable.]

The purpose of this chapter is to provide an understanding of the technical aspects of the data used for modeling, while also providing evidence that data weaknesses and limitations have been thoroughly researched and mitigated within the scope of the model development process.

The model's data sources and controls are discussed followed by an explanation as to how data is accessed, processed and centrally stored for all of the model development phases. Included is a description of the data reconciliation, preparation, and cleaning processes as well as any data sampling and transformations performed. Further, evidence of all business and technical data quality checks performed is included in this chapter. This chapter also includes a list of all relevant policy/operational events that impact the model results.

The key data enhancements attempted as part of the 2019 CCAR process includes the following-

1. Consistent product identification logic that align with the business definition and rendition

2. For 2019 CCAR model, one of the changes in the data source, specifically of the REL_TRAN.LOSS table has been the inclusion of CPB loan data. In coordination with Risk Data Mart and CPB Risk team a restatement of the loss table for all CPB historical data was implemented in Jan 2018. This restatement addresses the data challenges that were reported on the previous CCAR models, where CPB was manually included in the loss table provided by the CPB Risk team. This restatement automated the process, added data quality check, reconciliation between other data source and now available in a single loss table. See official email from Data Mart attached.

3. All source input files have been copied to a frozen dataset to prevent potential model errors from version updates

4. Extended the model development data to include recent time to holistically capture the portfolio composition changes

5. Incorporated balance differences between CMI and CPB

6. Introduced new interaction effects and splines to improve the model's performance in specific periods or segments

7. Imputed the missing values with a median, instead of omitting them and thereby reducing the possibilities of omitted variable bias

8. Capped the extremely large value of the continuous variable by 99th percentile plus 2 times the standard deviation

9. In PD model, for 2nd liens, instead of assuming a constant first lien balance when calculating its mark-to-market CLTV, assumed a constant junior ratio to impute the first lien balance and CLTV

10. Leveraged a new refreshed property valuation file, in line with what the business recommends and uses, for estimating the distressed values for all properties with/without prior distressed values

11. Included all zero losses in the development data pool for the Severity model

4.1 Model Input

This section will provide additional information to the "Automated Document" Section A.3.: "Inputs, Outputs & Assumptions" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

The data development phase, review phase and sampling phase are designed to determine whether building a risk model is feasible and to set high-level parameters such as loan exclusions, target definition, development sample window and performance sample window parameters. These high-level parameters decrease the number of anomalies and improve the accuracy of the model, and are naturally supported by a sound rationale. Significant time and engagement with stakeholders is the key to building a successful risk model throughout the data development, review, and sampling phases.

For 2019 Method A RM Model, model developers sourced data to build the model using both internal and external data sources. The data is compiled for use in model development and multiple quality checks are conducted to ensure continued accuracy of the data. The following sections describe the data sources and controls in further detail.

4.1.1 Data Sources and Controls

[Provide details regarding how the data was accessed, processed and stored.]

What are the performance, segmentation and account data sources for the model?

> [Document sources for all internal and external data used in model development. Describe each data source and whether the data is extracted in a manual or automatic manner. Mention whether the data has been used in the past for model development purposes.]

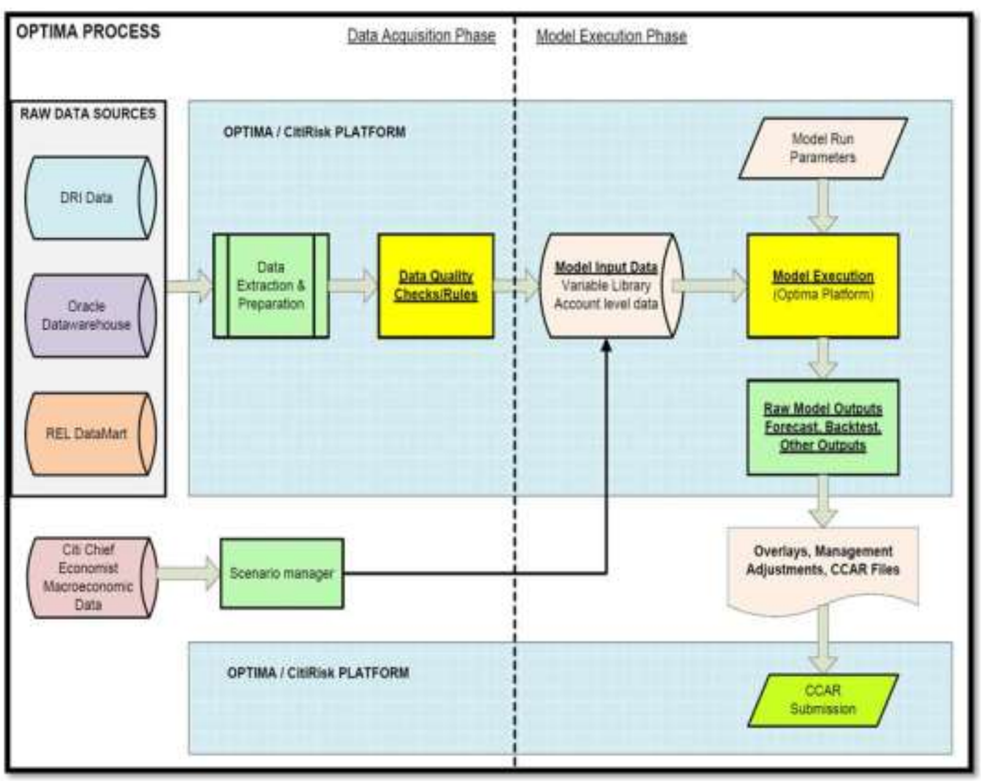
Model developers sourced data to build the model using both internal and external data sources. The data is compiled for use in model development and multiple quality checks are conducted to ensure accuracy of the data throughout the process. The following sections describe the data sources and quality control in further detail.

Significant time and engagement with stakeholders is key to building a successful risk model throughout the data development, review, and sampling phases. The process is designed to determine whether building a risk model is feasible and to set high-level parameters such as loan exclusions, target definition, development sample window, and performance sample window. These parameters decrease the number of anomalies while improving the accuracy of the model.

The 2019 Method A RM model development data and portfolio level data are all sourced (at the time of development) from robust internal data sources consisting of: SAS Grid/Optima/Model Execution Package and from an established external sources for macro-economic information (i.e. HPI, Unemployment, Interest Rate), Core Logic, Moody's, and Global Consumer Risk Management (GCRM).

As result of the new global initiatives within the Global Consumer group, the 2019 NA Mortgage Method A RM models leveraged this new centralized data source called MEP Optima (see figure 4.1.1.1). This Global Database platform within the SAS Grid environment which hosts multiple raw data sources as well as macroeconomic variables. Within the grid, a Model Execution Platform (MEP) contains all the data required to execute the models which by design have undergone several data quality check and data proofing process to make sure to comply all the data requirements for all CCAR models and data quality principles.

Figure 4.1.1.1 MEP Optima Process



The 2019 Method A RM model's objectives are to produce core line of business risk models with a product level focus on residential mortgage portfolio. This segmentation also aligns with the business strategy, and loss forecasting regulatory requirements. In addition, the data scope will be limited to BankMortgage Inc.(CMI), Bank Private Bank(CPB) and US Retail Bank (USRB) mortgage portfolios, including conventional mortgage loan products with non-zero balances, risk owned portfolio products.

MRM Question - Sponsor has mentioned creation of unique OOT sample is in response to limitation#19546. However, limitation #19546 states about inclusion of pre-2008 data in LGD model. Therefore, sponsor is requested to provide justification for not including pre-2008 data in LGD model development sample.

Answer - Although there are some reconciled issues caused by the misclassified loans mainly observed before May 2008, CAMU had set the observations date and performance window for development data from Feb 2006 to Dec 2017. The decision was made mainly to ensure there are sufficient observations to meet the gating principles outlined by Model Risk Management Policy, and to cover both stress and non-stress periods within the U.S. housing price lifecycle. However, abnormal actual losses were observed in snapshot Jan-2008 after removing these loans from the model back testing. To avoid any misinterpretation of the model performance, shifting the reporting snapshot date from Jan-2008 to April-2008 is considered appropriate. Please refer to Section 6.3 for the rationale of excluding 2008Q1 data from back testing.

Table 4.1.1.1 Data Source, Location, and Frequency of Updates

Data	Location and Source	Owner	Frequency Updated
Loan performance and borrower data	SAS Grid/Model Execution Package	REL Risk Data Team	Monthly
Bureau data	SAS Grid/Model Execution Package	REL Risk Data Team	Monthly
Loan Status/Resolution data	SAS Grid/Model Execution Package	O&T Operations	Monthly
Historical Macroeconomic data HPI Unemployment Interest Rate LIBOR Income VIX S&P500 index All Other Rates	Core Logic (Moody's) www.economy.com Bloomberg Haver Bank Velocity SAS Grid/Model Execution Package	External Vendor/REL Risk Data Mart/ GCRM/Risk Portfolio Loss Forecasting	Monthly
Future Macroeconomic data	Forecast provided by Bank's GCRM to a controlled access internal Share Point Site available for download (https://globalconsumer.collaborationtools.consumer.Bankgroup.net/sites/CCAR_RR/default.aspx)	GCRM/Risk Portfolio Loss Forecasting	Semi-annually

What are the macroeconomic data sources for the model including stress scenarios?

[Document sources for all internal and external data used in model development. Verify data source consistency across models developed in the same country or business – macroeconomic data for the same variable (e.g. GDP) should be identical.

Describe each data source and whether the data is extracted in a manual or automated manner. Only Bank-verified data sources are acceptable. Mention whether the data has been used in the past for model development purposes.

If there were alternative sources of macroeconomic data, provide rationale for selecting the specific source.]

Example: *Macro-economic scenarios are sourced from Federal Reserve CCAR guidance, in conjunction with Bank-verified data.*

There are two main categories for the macroeconomic data used in the model development and sensitivity analysis namely: Historical and Future Outlook macroeconomic data. GCRM provides both historical and future macroeconomic variables for US State level GDP, S&P 500 index, Volatility Index (VIX), Income and Interest rates (Swap 1 year, 5 year and 10 year, 30 year Prime rate etc.), which are sourced originally, from Bloomberg, Bank Velocity and Haver. For historical Home Price Index (HPI), Risk Data Mart team managed and own the processing of the data originally sourced from Core Logic. Historical Unemployment rate, Risk Portfolio Loss Forecasting Team processes the data from Moody's economy.com platform before loading to SAS Grid/Model Execution Package in Optima.

All future macroeconomic data are provided by GCRM and then process and managed by Risk Portfolio Loss Forecasting team leverage in BAU Loss Forecasting exercise and analysis. Both historical and future macroeconomic data undergo systematic control and data quality process before loading to SAS Grid/Model Execution Package in Optima for model development usage.

What process checks and controls are in place to ensure data integrity and that the data is properly sourced and reliable?

- What process checks and controls are in place to ensure data integrity and that the data is properly sourced and reliable?

  [Describe the high-level data preparation and transmission process, with a focus on maker-checker responsibilities, accountability of key personnel during the process, and data versioning. Describe how the model sponsor ensured that developers received current data (including the sign-off process, and when this occurs), and that this current data was used in model development. Mention sign-offs and other process checks confirming data adequacy for the modeling purpose within each process step.

  What data was considered to be potentially relevant but was not used? What is the sponsor's view of the potential impact of excluding this data? [Please provide the details and the impact analysis]

Describe whether the process differs for model implementation data (compared to model development data). Describe the process both for performance data, and macroeconomic data and stress scenarios.

This section also includes controls for data processing scripts. If a standard version/revision control system is used (such as SVN, CVS, Git, ClearCase or Team Foundation Server), describe the way in which versioning control is used. Describe controls in place to:

- Track changes made to data extraction and processing scripts.
- Track which version of the scripts was used to prepare a given version of the data.
- Prevent inadvertent changes in data extraction scripts.

Describe if a code review was performed (script maker/programmer - script checker split).]

Historical datasets stored in SAS Grid Model Execution Package (MEP) in Optima and Risk REL Data Mart, undergo an automated and controlled extraction, transformation and loading (ETL) process, which include detail data validation and data quality procedure for all variables and attributes. In addition to data quality checks by the Bank's MEP Optima team, Risk Management has quarterly Management Control Assessments (MCAs) setup to review risk associated with inadequate monitoring, improper data transformation, and improper changes to data, business policies, or business logic.

Additional layer of maker-checker control process is setup between GCRM, Risk Portfolio Loss Forecasting team and CAMU in handling Macroeconomic data. Please refer to attachment '4.1.1 GCRM Evidence_Macro_Inputs' for the GCRM evidence around the usage of all macro-economic variables in the models. CAMU Data Quality team performs additional validation check before using them for Model development. The maker-checker process comprised of checking data integrity, independent validation of the SAS code logic and proper signoff on any data treatment and adjustment made on the original data source.

Furthermore, CAMU Data Quality team performs detail reconciliation from the original data source and collaborates with Risk Data Team to identify gaps and document resolution. CAMU is member of the data steward's committee and work with other risk business partners to provide proper escalation and resolution on the all data issues. Moreover, CAMU production and implementation team prepare the data and provides clean and standardized dataset to modelers. Finally, modelers further evaluate the data by analyzing attributes, characteristics performance measures and key risk metrics to identify inherent potential data issue before using in model development.

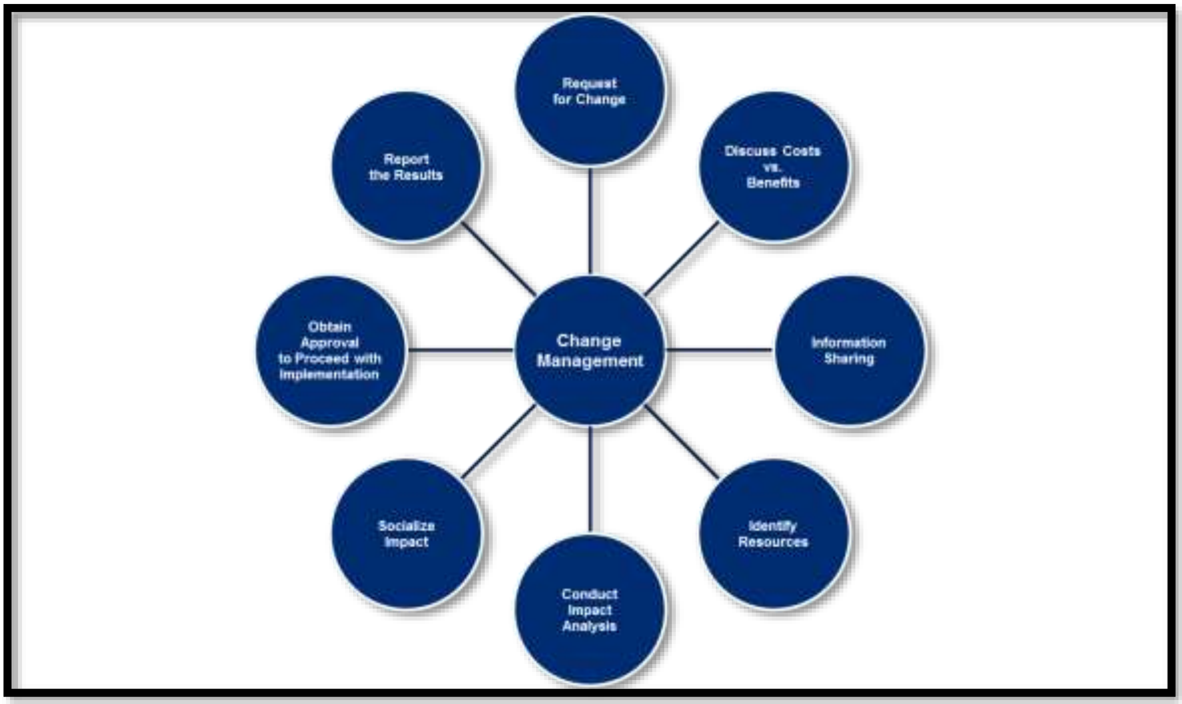
Table 4.1.1.2 MEP Optima Extract, Transform, and Load (ETL) Processing

MEP Optima Extract, Transform, and Load (ETL) Processing

Controls	Monitoring/MIS
1. Autosys enables the establishment of automated schedules and user defined event rules for the execution of required programs. Rules and prerequisites can be loaded into Autosys that must be satisfied prior to the execution of a program. Autosys provides real-time alerts and user notifications through-out the execution of the job schedule. Every job has an assigned owner, primary monitor, secondary monitor, business admin, technical admin, and subject matter expert. The MEP Optima Team performs/monitor these controls utilizing Autosys Scheduler, Job Metadata SAS datasets and SAS Logs/Reports on RELSAS20 Unix server providing information about each job's scheduling details, assignments, and expected outcomes	i. Direct Negative Control Information ii. Activity logs stored in SAS datasets on SAS Grid and Unix server iii. Attributes, characteristics, or conditions satisfied: job is executed successfully with a passing grade.
2. A file transfer reconciliation is performed from data source to the data copied. This reconciliation deals with mainly record count. An e-check is in place to notify Credit Risk when a trend occurs in record count that is above a certain percentage limitation.	i. Direct Negative Control Information ii. Activity logs stored in SAS datasets on SAS Grid and Unix server iii. Attributes, characteristics, or conditions satisfied: job is executed successfully with a passing grade.
3. Decisions reviewed during the Risk Data Stewards meeting are to be acknowledged and approved by Charter Members of the Data Stewards by returning an email with their approval when meeting minutes are distributed. Email approvals are obtained from the appropriate end users management team prior to releasing monthly MEP Optima globally.	i. Timely approval of MEP Optima changes/updates by appropriate approval authorities. ii. To make sure changes to the MEP Optima are being communicated via Data Stewards meetings. iii. Monthly release of Model Execution Platform (MEP) is being approved by appropriate resources. iv. Share Point Data Steward meeting minutes are up-to-date and available.

Further, there is significant data governance for the overall management of the availability, usability, integrity, and security of the MEP Optima platform. The below process flow clearly outlines the flow for data governance and data change management within the repository:

Figure 4.1.1.2 Data Governance Flow for MEP Optima



- Is the macroeconomic data source and data itself approved for use in Bank?

[Describe if the macroeconomic data is acceptable for use in Bank models (e.g. that a sufficient forecast period is available for the data, that the source is considered reliable, etc.)

Yes. All macroeconomic sources mentioned above are approved for use in Bank CCAR Mortgage Risk Models, both historical and future macroeconomic data sources either are directly from GCRM or processed by Risk Portfolio Loss Forecasting team with corresponding approval and signoff confirmation from GCRM. Furthermore, macroeconomic stress scenarios are provided uses the Federal Reserve CCAR guidance in conjunction with Bank verified data.

4.1.2 Data Reconciliation

[The following template can be used to document reconciliation and as a guide through the reconciliation process. If no data was used in the model development, please answer "N/A – no data used" for all questions in this section.]

Is the data source regularly reconciled to another independent data source for all key model variables?

[Respond "Yes" or "No". If "Yes", describe the independent (finance, controlling etc.) system, the reconciliation process and sign-off process. Include how often the reconciliation is performed. Attach evidence either in the appendix of this document or reconciliation templates. If "No", provide an explanation for the lack of reconciliation, along with compensating controls and other measures made to ensure data reliability.

All key model variables should be reconciled. For example, if the model calculates a PD rate based on the ratio of active and default accounts, then the number of defaults and the number of active accounts should be reconciled. Reconciling only GCL and NCL is insufficient.]

Yes. Bank reconciles data at minimum on a quarterly basis as evidenced within the reconciliation template.

Bank's Risk Analysis and Reporting Team reconcile its Monthly CCAR Fed Submission (14M report) to the Federal Reserves Consolidated Financial Statements for Holding Companies Report (FR Y-9C). There are some differences between the two reports which are due to varying accounting treatment of loans as reflected in the 14M and FRY-9C reporting, which are discussed in more detail below.

In addition, Optima group performs reconciliation and variable profiling for all root source variables uploaded into SAS Grid MEP in Optima. Detail data quality check are performed to make sure all the data elements are within the proper expected domain values and monitor any outliers or unexpected data issue encountered.

If the data source is not regularly reconciled, how was the data used for model development reconciled?

[Describe the independent (finance, controlling etc.) data system. It should be clear that the system is independent. It should be a system used by another department, preferably used to prepare financial statements.

Describe the reconciliation process. Attach evidence either in the appendix of this document or the reconciliation templates. Describe any differences in the process between development and implementation reconciliation.

If the model is prepared by an external model developer, or if there is a complex data handover, describe how the model developer verified that the data used for model development is complete and matches data extracted by country. The suggested approach is to obtain data reconciliation from the model sponsor and verify that the reconciliation matches data used for actual model development. A complex data handover takes place if the model is developed out of country, e.g., by a shared service center or Bank model development hub.

All key model variables must be reconciled. For example, if the model calculates a PD rate based on the ratio of active and default accounts, then the number of defaults and the number of active accounts must be reconciled. Reconciling only GCL and NCL is insufficient.]

Not applicable. See reconciliation evidence as provided in the required template.

Were there any reconciliation differences?

[Describe any reconciliation differences encountered during the process, high-level causes of the differences, and suggested remediation. It must be clear whether the data can or cannot be used for model development, and whether it is sufficiently accurate. Quantify the impact of differences.]

The first difference is that the 14M report leverages UPB (unpaid balances) while the FR Y-9C Financial Report leverages ENR (ending net receivables). UPB calculations do not allow for movement of a loan from one portfolio to another. Although the mortgage portfolios do not have a large number of internal portfolio transfers on a regular basis, there is a segment of HELOC loans (which at any given month are converting to fully amortized loans) which can cause some insignificant discrepancies quarter-over-quarter.

Secondly, CAMU would like to point out these distinct areas of insignificant differences: 1) FAS91 (Accounting for Nonrefundable Fees and Costs Associated with Originating or Acquiring Loans and Initial Direct Costs of Leases), 2) LHS Clearing loans (loans in process pipeline) and 3 Offline FAS91 (HELOC Mods)

Therefore, in the enclosed reconciliation template, the spreadsheet attempts to reconcile these differences within a tolerance of less than 1%.

In addition to the overall ENR balance reconciliation between Risk system and Financial General Ledger, individual key attributes were also tested for consistency and accuracy by going through the development sample attributes and compared against the original data source from the SAS Grid/MEP Optima over a snapshot period consistent in the development time frame; overall results are consistent with the official data source, with very minor volume of difference due to timing or re-statement of the original data source system. No serious pattern or data quality anomalies were observed.

The reconciliation process includes frequency matching using a sampling methodology from the development sample used in the modeling process with specific focus on the data quality and consistency of the data used in the model development and original data source.

Remaining variances are mainly due to a timing difference between the time the modeler used the dataset and the time the validation check was performed. The DataMart team is continuously improving and updating the source data file based on the latest and most accurate information. This may result in historical data adjustments or re-statements. (See attached files named ‘4.1.2 CCAR Data Reconciliation…’ for detail results).

4.1.3 Data Characteristics, Fields and Definitions

[Provide a description of data preparation, including the steps followed in assembling the data, and the number of records (observations) per account. Consider including stylized or actual example records in order to make the structure of the data set transparent to the reviewer. Describe each input explicitly, including dependent/independent variables (for regressions), data sources, frequency of data pull, contact person, format (e.g. Boolean, character, integer, floating-point number, fixed-point number and length, size of the variable and position of the decimal point), proper unit (e.g. thousands, millions, percentage), data type (external, internal), derived variables construction methodology, and output. For monetary variables, indicate whether USD or another currency is used and label all graphs and tables accordingly.]

Please provide descriptive statistics for the underlying data including histograms/frequency tables, and trends over time.

[Please provide mean, standard deviation, median, extreme values, mode, percentage of missing records, and other statistical descriptive measures (as appropriate), for each input.

Comment on whether the values are in line with expectations given the composition of the portfolio. Analyze outliers.]

PD Model

The initial monthly data for Method A RM PD model is from Jan 2005 through Dec 2017 yielded 77,969,449 records with 4,785,882 unique accounts before data filtering in Table 4.1.3.1. The data is constructed as a standard panel dataset and the performance of the monthly data is followed month over month until the onset of one of these statuses: involuntary payoff (IVP), or voluntary payoff (PO), or survived at the end of Dec 2017.

After data is extracted from the risk database, steps are taken to remove and/or clean erroneous data. Any lines with a data quality issue are excluded from the data water fall. In addition to the data clean up, CAMU determined to exclude all observations for loans received treatment since there are models designed particularly for this segment. CAMU also collaborated with the risk team to ensure the correct population is included in the development data. The detail of all data quality issues leading to exclusions are noted in the data waterfall in 4.1.3.1.

The exclusions were generally very limited and designed in alignment with business knowledge/intuition. The criteria can be categorized into three exclusion reasons.

Exclusion Reason 1: Criterion # 1- These are excluded in order to define the model development population that best aligns with CCAR objectives. Particularly the asset sale Sold and Released (SR) or Sold and Serviced (SS) records in the performance months are excluded, for the reason that SR do not have reliable performance, while SS is not deemed as portfolio risk hence considered outside CCAR scope. Indeed, the drop of 14,305 unique loans at this step is primarily caused by excluding sold and released – they only have two performance records, first in the observation month is portfolio loan, and second in the performance month when the asset becomes Sold and Service Released.

Exclusion Reason 2: Criterion #s 2, 4,5 and 6 -These are excluded to improve data quality and remove the records with missing performance or negative balances or those outside the scope of model development period.

Exclusion Reason 3: Criterion # 3 - When a loan is successfully modified, its delinquency level will in general be set to current. This is not due to borrower's self-curing behavior and should not be modeled as a curing event. The Risk Loss Mitigation Policy instead of borrower behavior primarily determined the volume of modifications. The modified pools of loans are modeled within the scope of the Modified loan model suite and hence are excluded from the RM development data.

Table 4.1.3.1 Data Water Fall for Method A RM Model

Exclusion Detail		Remain. # of Records	# of Records Exclud.	% of Records Excluded	Remain. # of Unique Loan	# of Unique Loan Exclud.	% Unique Loan Exclud.
RM Portfolio (CMI &CPB)		77,969,499			4,785,882		
1	Exclude loans sevice released in performance month	77,955,194	14,305	0.0%	4,784,752	1,130	0%
2	Exclude loans with termination date before observation month, or unknown/unclear status in performance month	77,804,196	150,998	0.2%	4,782,042	2,710	0%
3	Exclude loans with unknown/unclear status at performance month, or modified in performance month	77,499,931	304,265	0.4%	4,764,547	17,495	0%
4	Exclude loans with missing or zero UPB in observation month	77,088,784	411,147	0.5%	4,736,849	27,698	1%
5	Exclude loans terminated but with unclear termination status in performance month	77,083,582	5,202	0.0%	4,736,432	417	0%
6	Exlude loans with negative balance in observation month	77,083,574	8	0.0%	4,736,432	0	0%
Total		77,083,574	885,925	1.1%	4,736,432	49,450	1.0%

From Jan 2005 to Dec 2017, there are a total of 4,785,882 unique accounts with 77,969,499observations pulled from the monthly Residential Mortgage portfolio of which 75,514,848 observations are CMI and 2,454,651 are from CPB first Lien. After the data waterfall, the data contains 4,736,432 unique accounts with a total of 77,083,574 observations.

In the above panel data, depending on the start buckets, five major scenarios can typically describe a borrower's behavior that include (1) unchanged delinquency status, (2) cure or partial cure status: from a higher delinquency status to buk01 or to a relative lower delinquency status, (3) status goes worse: from a lower delinquency status to a higher delinquency status, (4) involuntary payoff, or (5) voluntary payoff.

The 7 starting buckets are defined as the delinquency state at the observation month:

1. Buk 01: 0 to 29 days past due

2. Buk 2: 30 to 59 days past due
3. Buk 3: 60 to 89 days past due
4. Buk 4: 90 to 119 days past due
5. Buk 5: 120 to 149 days past due
6. Buk 6: 150 to 179 days past due
7. Buk 7: at least 180 days past due

The model development sample is then segmented in Table 4.1.3.2 by delinquency status that can be represented by a 7X9 transition matrix in which loans transitioned from one of the seven delinquency states in the observation month (t) to the nine possible transition states including the terminal events of charge off (IVP) and payoff (VP) in the performance month (t+1).

In Table 4.1.3.2 below, all cells highlighted in green are the transitions that will be estimated as a model with key risk drivers (modeled transition), and the cells in pink are being used as reference category when develop the model. In addition, the cells in white are the non-modeled transitions to be developed as a lookup table with constant numbers for stress and non-stress period respectively.

Table 4.1.3.2 Transition Roll Rate Matrix of RM panel data before Sampling on Current Bucket

Before Sampling (Observation month 01/2005–12/2017)										
status in observation month	status in performance month									Total
	BUK01 (0–29 DPD)	BUK2 (30–59 DPD)	BUK3 (60–89 DPD)	BUK4 (90–119 DPD)	BUK5 (120–149 DPD)	BUK6 (150–179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0–29 DPD)	65,475,905 96.7%	1,401,543 2.1%	20,175 0.0%	4,976 0.0%	1,740 0.0%	775 0.0%	5,111 0.0%	1,823 0.0%	807,174 1.2%	67,719,222
BUK2 (30–59 DPD)	930,611 30.2%	1,428,806 46.3%	666,285 21.6%	11,122 0.4%	2,797 0.1%	678 0.0%	4,391 0.1%	870 0.0%	39,369 1.3%	3,084,929
BUK3 (60–89 DPD)	132,840 11.4%	184,152 15.9%	393,540 33.9%	428,592 36.9%	5,458 0.5%	1,363 0.1%	4,182 0.4%	1,141 0.1%	10,285 0.9%	1,161,553
BUK4 (90–119 DPD)	39,220 5.9%	24,110 3.7%	63,089 9.6%	167,235 25.4%	351,279 53.3%	3,017 0.5%	4,005 0.6%	2,453 0.4%	4,821 0.7%	659,229
BUK5 (120–149 DPD)	20,039 3.9%	5,598 1.1%	9,513 1.9%	29,499 5.8%	87,716 17.2%	344,003 67.5%	4,673 0.9%	5,264 1.0%	3,420 0.7%	509,725
BUK6 (150–179 DPD)	14,753 3.5%	2,967 0.7%	2,808 0.7%	5,751 1.3%	18,294 4.3%	60,024 14.1%	287,351 67.3%	32,060 7.5%	3,124 0.7%	427,132
BUK7 (180+ DPD)	52,069 1.5%	7,537 0.2%	3,699 0.1%	4,411 0.1%	7,318 0.2%	17,774 0.5%	3,257,388 92.5%	143,930 4.1%	27,658 0.8%	3,521,784
Total	66,665,437	3,054,713	1,159,109	651,586	474,602	427,634	3,567,101	187,541	896,851	77,083,574

CAMU conducted random sampling on the population that stays in BUK01 in both the observation month and performance month with a 2% sampling rate. As shown in table 4.1.3.3, number of records in the cell of "BUK01-BUK01" decreased from 65,475,905 to 1,307,971. After sample down, the total number of records drops to 12,915,640 of which 12,794,442 are CMI and 121,198 from CPB first lien. Table 4.1.3.4 gives the records and the transition rate on the sampled data with weight, which shows that the randomly sampling significantly reduces number of records to be included in modeling while maintaining the same transition rate.

Table 4.1.3.3 Number of Records in RM data after 2% Sampling on Current Bucket

After Sampling without weight (Observation month 01/2005-12/2017)										
status in observation month	status in performance month									Total
	BUK01 (0-29 DPD)	BUK2 (30-59 DPD)	BUK3 (60-89 DPD)	BUK4 (90-119 DPD)	BUK5 (120-149 DPD)	BUK6 (150-179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0-29 DPD)	1,307,971	1,401,543	20,175	4,976	1,740	775	5,111	1,823	807,174	3,551,288
BUK2 (30-59 DPD)	930,611	1,428,806	666,285	11,122	2,797	678	4,391	870	39,369	3,084,929
BUK3 (60-89 DPD)	132,840	184,152	393,540	428,592	5,458	1,363	4,182	1,141	10,285	1,161,553
BUK4 (90-119 DPD)	39,220	24,110	63,089	167,236	351,279	3,017	4,005	2,453	4,821	659,229
BUK5 (120-149 DPD)	20,039	5,598	9,513	29,499	87,716	344,003	4,673	5,264	3,420	509,725
BUK6 (150-179 DPD)	14,753	2,967	2,808	5,751	18,294	60,024	287,351	32,060	3,124	427,132
BUK7 (180+ DPD)	52,069	7,537	3,699	4,411	7,318	17,774	3,257,388	143,930	27,658	3,521,784
Total	2,497,503	3,054,713	1,159,109	651,586	474,602	427,634	3,567,101	187,541	895,851	12,915,640

Table 4.1.3.4 Transition Roll Rate Matrix of RM panel data after Sampling on Current Bucket with weight

After Sampling with weight (Observation month 01/2005~12/2017)										
status in observation month	status in performance month									Total
	BUK01 (0~29 DPD)	BUK2 (30~59 DPD)	BUK3 (60~89 DPD)	BUK4 (90~119 DPD)	BUK5 (120~149 DPD)	BUK6 (150~179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0~29 DPD)	65,398,550 96.7%	1,401,543 2.1%	20,175 0.0%	4,976 0.0%	1,740 0.0%	775 0.0%	5,111 0.0%	1,823 0.0%	807,174 1.2%	67,641,867
BUK2 (30~59 DPD)	930,611 30.2%	1,428,806 46.3%	666,285 21.6%	11,122 0.4%	2,797 0.1%	678 0.0%	4,391 0.1%	870 0.0%	39,369 1.3%	3,084,929
BUK3 (60~89 DPD)	132,840 11.4%	184,152 15.9%	393,540 33.9%	428,592 36.9%	5,458 0.5%	1,363 0.1%	4,182 0.4%	1,141 0.1%	10,285 0.9%	1,161,563
BUK4 (90~119 DPD)	39,220 5.9%	24,110 3.7%	63,089 9.6%	167,235 25.4%	351,279 53.3%	3,017 0.5%	4,005 0.6%	2,453 0.4%	4,821 0.7%	659,229
BUK5 (120~149 DPD)	20,039 3.9%	5,598 1.1%	9,513 1.9%	29,499 5.8%	87,716 17.2%	344,003 67.5%	4,673 0.9%	5,264 1.0%	3,420 0.7%	509,725
BUK6 (150~179 DPD)	14,753 3.5%	2,967 0.7%	2,808 0.7%	5,751 1.3%	18,294 4.3%	60,024 14.1%	287,351 67.3%	32,060 7.5%	3,124 0.7%	427,132
BUK7 (180+ DPD)	52,069 1.5%	7,537 0.2%	3,699 0.1%	4,411 0.1%	7,318 0.2%	17,774 0.5%	3,257,388 92.5%	143,930 4.1%	27,658 0.8%	3,521,784
Total	66,588,082	3,054,713	1,159,109	651,586	474,602	427,634	3,567,101	187,541	895,851	77,006,219

At the model development stage, it is further determined to exclude the records collected on observation month before Feb 2006 because of the relatively low data quality of the credit bureau attributes such as FICO during that period. It's also decided that the model to be developed only on loans within the range of CMI or CPB official loan list. The official loan list is the list that Risk Loss Forecasting team uses to report their number for CCAR. Official loan list is reconciled with PEARL system and also reconciled in Optima as part of the data quality check. This is the main reason the CCAR model is developed on the official loan list to make sure it is based on reconciled data which is reported by regulators, thus showing correct and consistent population for model development.

Hence 1,727,861 more records were removed from the sampled data with 11,187,779 observations left for model development. By applying weight on the BUK01->BUK01 sample, the total number of observation is 63,657,028.

All records that have gone through the above waterfall and sampling process are them divided into model development (DEV), in-time hold out (INT) validation and out-of-time (OOT) validation sample. First, data collected between Feb 2006 and Dec 2011, and Apr 2014 and Dec 2017 are used as the in time sample, and those between Jan 2012 and Mar 2014 are kept as the out-of-time validation sample.

In this version of model development, CAMU holds out an OOT sample solely to comply with MRM's OOT test requirement although CAMU believes that holding out OOT data will potentially result in loss of performance and portfolio change information. Due to this concern, CAMU carefully chose Jan-2012 to March-2014 period as the out-of-time validation sample and retained data between Feb 2006 - Dec 2011 and Apr 2014 -Dec 2017 as model development data due to the following considerations –

1. First and foremost, as per limitation # 19546, it was recommended to include the most recent time period as part of the model development sample. To comply with MRM's requirements, CAMU did not utilize the hold out most recent data as OOT. CAMU also agreed that holding out recent performance

data would disable us from capturing the most recent macro-economic trend, such as interest rate increase, portfolio and underwriting policy change. The inclusion of the recent period data helped capture the go-forward state of the business in terms of origination profile and portfolio composition mix.

2. Excluding 2006-2008 data is not an option because it will result in the loss of a portion of the stress period performance. This leaves CAMU no choice but to carve out a middle period as the OOT hold out sample to follow MRM's OOT requirement.

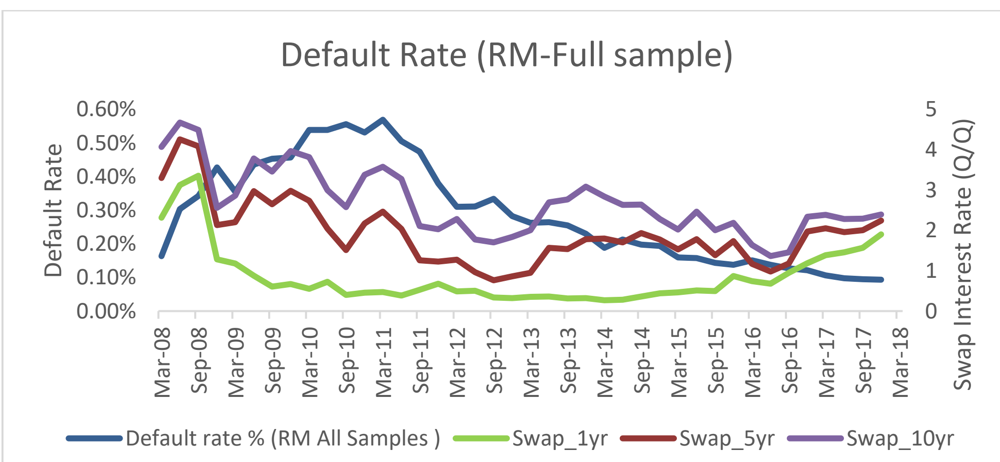
3. To ensure parameter stability and OOT hold out didn't cause significant information loss, CAMU conducted further parameter stability analysis and revised/dropped instable parameters. The parameter stability analysis was two-fold including:

   a. CAMU estimated the same specification on the 20% INT sample to ensure consistent parameter sign, significance and reasonable magnitude

   b. CAMU further estimated the same specification on the entire data including the OOT period to ensure consistent parameter sign, significance and reasonable magnitude

Figure 4.1.3.1a demonstrates how the default rate in the development data captures the change of swap interest rates during the stress and non-stress periods. When swap rates dropped significantly in 2009 and remained low between 2009 and 2011, higher default rate was observed. Likewise, when the interest rates increased slightly after the US economy stress period, the default rate continued to drop over time.

Figure 4.1.3.1a Residential Mortgage Development Default Rate and Swap Trends



Second, within the in-time sample, 80% of unique accounts with 43,645,214 observations were used as development samples vs 20% of unique accounts with 10,915,090 observations as validation. Table

4.1.3.5 givens the total number of records in the final model DEV, INT and OOT sample on the sampled data before and after applying weight.

It is worth noting though, that as CAMU expected, holding out the OOT data could potentially result in loss of performance information, especially given that the Jan2012-March2014 period represented the initial phase of economic recovery with high volume of various types of loan modifications and refinancing due to extremely low interest rate and foreclosure settlement (2012) etc. However, given the model's strong performance (backtest) for the OOT period, it can be concluded that the impact from such exclusion is expected to be small as the model development data still comprised of sufficient performance data covering both economic boom and bust. The model also included a full interest rate cycle in its development data set which included periods of rising and falling rates. Please refer to Chapter 2 for additional details on the interest rate cycle. A few transitions are heavily impacted by portfolio management policy over time, such as IVP transitions, are always statistically challenging to model in the first place. For additional details on the OOT backtest, please refer to Section 6.3 of the MDD. Please note the OOT sample only excluded model performance data specific to that period, loans that originated during the OOT period and were active post the conclusion of the OOT period are included as part of the development data sample.

Table 4.1.3.5 Sample size in Development/In-time Validation/Out-of-Time Validation Sample

data_use	Observation window	Frequency w/o weight	Frequency w/ weight
Development	Feb 2006~Dec 2011 & Apr 2014~Dec 2017 (80%)	7,557,694	43,645,214
In-time Validation	Feb 2006~Dec 2011 & Apr 2014~Dec 2017 (20%)	1,887,526	10,915,090
Out-of-time Validation	Jan 2012~Mar 2014	1,742,559	9,096,724

Table 4.1.3.6.1-3 provides the final one-month-ahead transition matrix of the model development, in-time-hold-out and out-of-time validation samples that are used estimate the CCAR 2019 never-modified RM PD model.

Table 4.1.3.6.1 Transition matrix of RM Development Sample

One-month-ahead Transition Matrix of RM model DEV sample (80% of 02/2006~12/2011 and 04/2014~12/2017)										
status in observation month	status in performance month									Total
	BUK01 (0-29 DPD)	BUK2 (30-59 DPD)	BUK3 (60-89 DPD)	BUK4 (90-119 DPD)	BUK5 (120-149 DPD)	BUK6 (150-179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0-29 DPD)	36,824,000 96.6%	842,319 2.2%	11,804 0.0%	2,918 0.0%	1,015 0.0%	420 0.0%	2,704 0.0%	1,068 0.0%	436,266 1.1%	38,122,514
BUK2 (30-59 DPD)	547,150 29.8%	842,667 45.9%	412,634 22.5%	6,934 0.4%	1,751 0.1%	367 0.0%	2,329 0.1%	496 0.0%	20,348 1.1%	1,834,676
BUK3 (60-89 DPD)	78,463 11.2%	108,804 15.5%	232,004 33.0%	271,305 38.6%	3,569 0.5%	853 0.1%	2,368 0.3%	656 0.1%	5,405 0.8%	703,416
BUK4 (90-119 DPD)	23,370 5.7%	13,980 3.4%	37,985 9.3%	100,863 24.6%	224,584 54.9%	2,000 0.5%	2,379 0.6%	1,436 0.4%	2,693 0.7%	409,290
BUK5 (120-149 DPD)	11,894 3.7%	3,374 1.1%	5,675 1.8%	17,883 5.6%	53,386 16.6%	220,750 68.8%	2,833 0.9%	3,304 1.0%	1,829 0.6%	320,928
BUK6 (150-179 DPD)	8,727 3.2%	1,765 0.7%	1,718 0.6%	3,588 1.3%	11,187 4.1%	36,578 13.6%	182,199 67.5%	22,266 8.3%	1,820 0.7%	269,848
BUK7 (180+ DPD)	29,406 1.5%	4,488 0.2%	2,252 0.1%	2,776 0.1%	4,484 0.2%	10,801 0.5%	1,826,620 92.0%	88,323 4.5%	15,392 0.8%	1,984,542
Total	37,523,010	1,817,397	704,072	406,267	299,966	271,769	2,021,432	117,548	483,753	43,645,214

Table 4.1.3.6.2 Transition matrix of RM In-Time Validation Sample

One-month-ahead Transition Matrix of RM model INT sample (20% of 02/2006~12/2011 and 04/2014~12/2017)										
status in observation month	status in performance month									Total
	BUK01 (0-29 DPD)	BUK2 (30-59 DPD)	BUK3 (60-89 DPD)	BUK4 (90-119 DPD)	BUK5 (120-149 DPD)	BUK6 (150-179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0-29 DPD)	9,211,800 96.6%	210,530 2.2%	2,924 0.0%	743 0.0%	239 0.0%	107 0.0%	612 0.0%	246 0.0%	109,065 1.1%	9,536,266
BUK2 (30-59 DPD)	136,268 29.8%	209,975 45.9%	103,227 22.6%	1,778 0.4%	413 0.1%	87 0.0%	567 0.1%	115 0.0%	5,029 1.1%	457,459
BUK3 (60-89 DPD)	19,983 11.4%	26,775 15.3%	56,932 32.5%	68,125 38.9%	911 0.5%	195 0.1%	555 0.3%	161 0.1%	1,375 0.8%	174,932
BUK4 (90-119 DPD)	5,777 5.6%	3,444 3.4%	9,530 9.3%	24,845 24.3%	56,641 55.4%	498 0.5%	551 0.5%	368 0.4%	634 0.6%	102,288
BUK5 (120-149 DPD)	2,987 3.7%	814 1.0%	1,476 1.8%	4,512 5.6%	13,281 16.5%	55,682 69.0%	606 0.8%	810 1.0%	486 0.6%	80,734
BUK6 (150-179 DPD)	2,261 3.3%	473 0.7%	459 0.7%	873 1.3%	2,829 4.2%	9,131 13.5%	45,842 67.5%	5,558 8.2%	443 0.7%	67,869
BUK7 (180+ DPD)	7,546 1.5%	1,225 0.2%	518 0.1%	691 0.1%	1,096 0.2%	2,682 0.5%	456,100 92.0%	21,775 4.4%	3,909 0.8%	495,542
Total	9,386,542	453,236	175,066	101,567	75,410	68,382	504,913	29,033	120,941	10,915,090

Table 4.1.3.6.3 Transition matrix of RM Out-of-time Validation Sample

One-month-ahead Transition Matrix of RM modelOOT sample (01/2012~03/2014)										
status in observation month	status in performance month									Total
	BUK01 (0-29 DPD)	BUK2 (30-59 DPD)	BUK3 (60-89 DPD)	BUK4 (90-119 DPD)	BUK5 (120-149 DPD)	BUK6 (150-179 DPD)	BUK7 (180+ DPD)	IVP	VP	
BUK01 (0-29 DPD)	7,504,250 97.4%	123,427 1.6%	1,471 0.0%	266 0.0%	109 0.0%	55 0.0%	550 0.0%	303 0.0%	74,445 1.0%	7,704,876
BUK2 (30-59 DPD)	85,581 25.3%	180,601 53.4%	68,162 20.2%	690 0.2%	128 0.0%	55 0.0%	428 0.1%	111 0.0%	2,331 0.7%	337,977
BUK3 (60-89 DPD)	12,548 8.4%	23,312 15.5%	67,026 44.6%	45,503 30.3%	369 0.2%	114 0.1%	444 0.3%	132 0.1%	825 0.5%	150,262
BUK4 (90-119 DPD)	3,518 4.4%	3,170 3.9%	8,866 11.0%	27,218 33.7%	36,445 45.2%	271 0.3%	459 0.6%	289 0.4%	438 0.5%	80,674
BUK5 (120-149 DPD)	1,584 2.7%	605 1.0%	1,157 2.0%	3,559 6.0%	12,618 21.3%	38,293 64.8%	654 1.1%	341 0.6%	318 0.5%	59,129
BUK6 (150-179 DPD)	1,318 2.7%	328 0.7%	334 0.7%	586 1.2%	1,867 3.8%	7,669 15.6%	34,080 69.2%	2,864 5.8%	231 0.5%	49,267
BUK7 (180+ DPD)	6,432 0.9%	896 0.1%	515 0.1%	486 0.1%	846 0.1%	1,822 0.3%	681,091 95.3%	21,010 2.9%	1,441 0.2%	714,539
Total	7,615,231	332,339	147,520	78,208	52,362	48,279	717,706	25,050	80,029	9,096,724

LGD Model

All samples in the severity dataset experienced involuntary payoff (IVP) resulting in terminating by default between Feb2005 -June2017. An IVP event is identified primarily by one of the following statuses: Short Sale (SS), Deed in Lieu of Foreclosure (DL), Foreclosure – Third Party Sale (3F), REO (RO), Charge Off (CO), or REO Sold (RS) in the Inactive Detail file located in CCR Risk MEP Optima; or having foreclosure/ short sale activities in foreclosure related tables in DRI. To ensure the correct population used in the severity model, CAMU collaborated with the risk team and only selected loans as the development samples if they satisfied the data water fall filtering in Table 4.1.3.7. As showed in below table, there were 582,301 unique accounts that defaulted from February 2005 to June 2017. After the data waterfall, there were 304,926 accounts remained for development, validation, and out-of-time checking.

Table 4.1.3.7.: Data Water Fall for Severity Model

	Exclusion Detail	Remain. # Unique Line	#Unique Records Exclude.	% Unique Records Exclude.
	Loans with IVP terminal event as of Feb 2018	**582,301**	-	-
1	Excluded Loans with inactive date not in the preliminary development period: Feb2005~Jun2017	**575,413**	6,888	1.2%
2	Excluded non-risk loans	**426,932**	148,481	25.5%
3	Excluded repurchase loans	**421,710**	5,222	0.9%
4	Excluded government incentive loans	**416,843**	4,867	0.8%
5	Excluded loans with modification after inactive	**415,319**	1,524	0.3%
6	Excluded non-VA government loans	**355,052**	60,267	10.3%
7	Excluded loans without loss	**346,215**	8,837	1.5%
8	Excluded loans with missing original balance or property amount	**343,998**	2,217	0.4%
9	Excluded loans corresponding to one-timer loss adjustment[1]	**337,994**	6,004	1.0%
10	Further excluded loans before Jan 2008 to improve data quality	**304,926**	33,068	5.7%
	Total	304,926	277,375	47.6%

For additional information on one-timer events and treatments, please refer to attachment "3.2_Memorandum_One-Timer Exclusions - 111417.docx"

The NCL Severity model data segmented by three different loan types: (1) First Lien, (2) Second Lien, and (3) VA Loan. Except VA loans which are VA-insured, all models were designed for portfolios with loans not insured by government agency. Presented below in Tables 4.1.3.8 -4.1.3.11 illustrates the defaulted loan distribution by Modified vs Non-Modified loans, by product types (Residential vs HELOC vs FRHEL) and by model development and validation samples. Please note within the scope of the Method A RM Model suite, we consider the non-modified residential mortgages only.

Table 4.1.3.8: Summary of Defaulted Loan Distribution by MOD vs Non-MOD

Default of Residential by MOD status			
Portfolio	**non MOD**	**MOD**	**Total**
1st Lien	65,055	6,072	71,127
1.Residential Mortgage	64,817	6,063	70,880
4.CPB/PBG	238	9	247
2nd Lien	21,902	504	22,406
1.Residential Mortgage	21,902	504	22,406
VA	7,320	1,251	8,571
1.Residential Mortgage	7,320	1,251	8,571
Total	94,277	7,827	102,104

Table 4.1.3.9: Summary of Defaulted Loan Distribution by Product Type

Portfolio	1.RESIDENTIAL	2.HELOC	3.FRHEL	Total
1st Lien	**71,127**	**1,225**	**1,236**	**73,588**
1.Residential Mortgage	70,880			70,880
2.Home Equity		1,157	1,200	2,357
3.CBNA/USCCB/USRB		54	36	90
4.CPB/PBG	247	14		261
2nd Lien	**22,406**	**58,392**	**141,969**	**222,767**
1.Residential Mortgage	22,406			22,406
2.Home Equity		56,665	140,743	197,408
3.CBNA/USCCB/USRB		1,685	1,214	2,899
4.CPB/PBG		42	12	54
VA	**8,571**			**8,571**
1.Residential Mortgage	8,571			8,571
Total	**102,104**	**59,617**	**143,205**	**304,926**

Table 4.1.3.10: Summary of Defaulted Loan Distribution(Non-CPB) by Development & Validation Samples

Non-CPB	DEV	IN_VAL	OOT	Total
1st Lien	44,296	10,977	9,544	64,817
full	8,836	2,202	2,319	13,357
other	34,097	8,428	6,630	49,155
zero	1,363	347	595	2,305
2nd Lien	17,315	4,350	237	21,902
full	14,252	3,555	129	17,936
other	2,460	656	92	3,208
zero	603	139	16	758
VA	4,230	1,077	2,013	7,320
full	30	7	5	42
other	2,026	486	900	3,412
zero	2,174	584	1,108	3,866
Total	65,841	16,404	11,794	94,039

Table 4.1.3.11: Summary of Defaulted Loan Distribution(CPB) by Development & Validation Samples

CPB	DEV	IN_VAL	OOT	Total
full	11	7	6	24
other	101	29	79	209
zero	4	0	1	5
Total	116	36	86	238

After the additional filtering by modification status and product type, the final sample size for the First Lien segment was approximated around 64,817 of which 68% (44,296) was reserved for model development. Tables 4.1.3.10 and 4.1.3.11 illustrates the distribution of defaults with the in time samples. The next set of Figures and Tables illustrate the loss trend and loss type by lien position and VA for the model development sample. For additional details, please refer to attachment '4.1.3 –Severity Trends'.xlsx'.

Table 4.1.3.12: Distribution of Defaults by Loss Types (First Lien) for In Time Sample

1st Lien Default year	loss_ind			Total
	full	other	zero	
2008	1,688	10,709	225	12,622
2009	2,956	11,384	272	14,612
2010	2,467	10,122	476	13,065
2011	2,849	7,462	284	10,595
2012	1,522	4,065	273	5,860
2013	683	2,263	274	3,220
2014	501	1,756	190	2,447
2015	437	989	151	1,577
2016	175	431	94	700
2017	103	183	71	357
Total	13,381	49,364	2,310	65,055

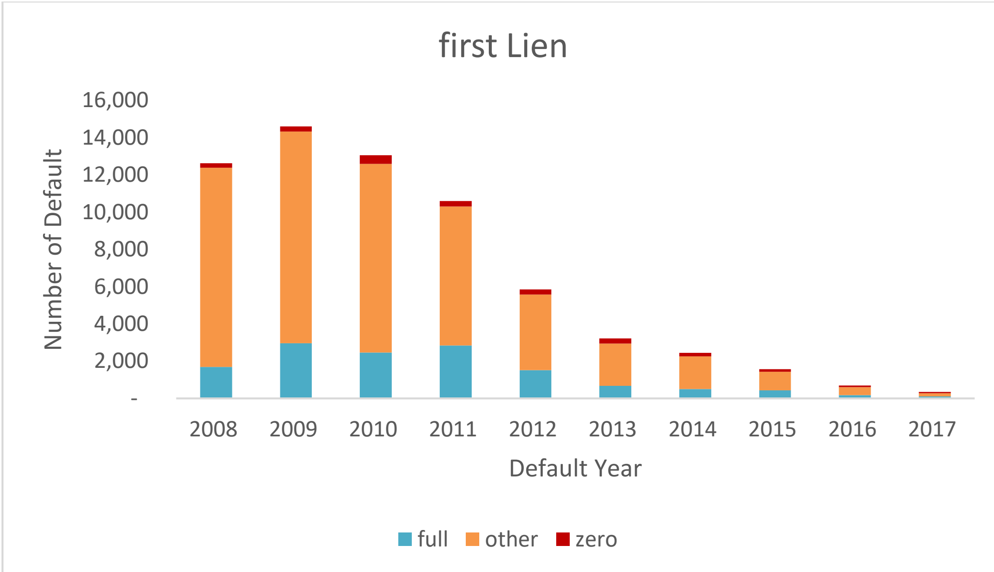
Table 4.1.3.13: Distribution of Defaults by Loss Types (Second Lien) for In Time Sample

2nd Lien	loss_ind			
Default year	full	other	zero	Total
2008	5,940	738	230	6,908
2009	6,307	1,056	368	7,731
2010	4,017	910	92	5,019
2011	1,523	391	48	1,962
2012	99	65	14	178
2013	26	21	2	49
2014	9	19	2	30
2015	8	6	2	16
2016	7	2	0	9
2017	0	0	0	0
Total	17,936	3,208	758	21,902

Table 4.1.3.13: Distribution of Defaults by Loss Types (VA) for In Time Sample

VA	loss_ind			
Default year	full	other	zero	Total
2008	14	121	67	202
2009	13	259	368	640
2010	2	754	694	1,450
2011	0	304	752	1,056
2012	4	367	496	867
2013	0	391	489	880
2014	2	574	431	1,007
2015	4	386	293	683
2016	2	188	204	394
2017	1	68	72	141
Total	42	3,412	3,866	7,320

Figure 4.1.3.1: Count of First Lien Defaults by Loss Types between Jan-2008 and June -2017.



first Lien
Number of Default
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
-
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Default Year
full
other
zero

Figure 4.1.3.2: Count of Second Lien Defaults by Loss Types between Jan-2008 and June -2017.

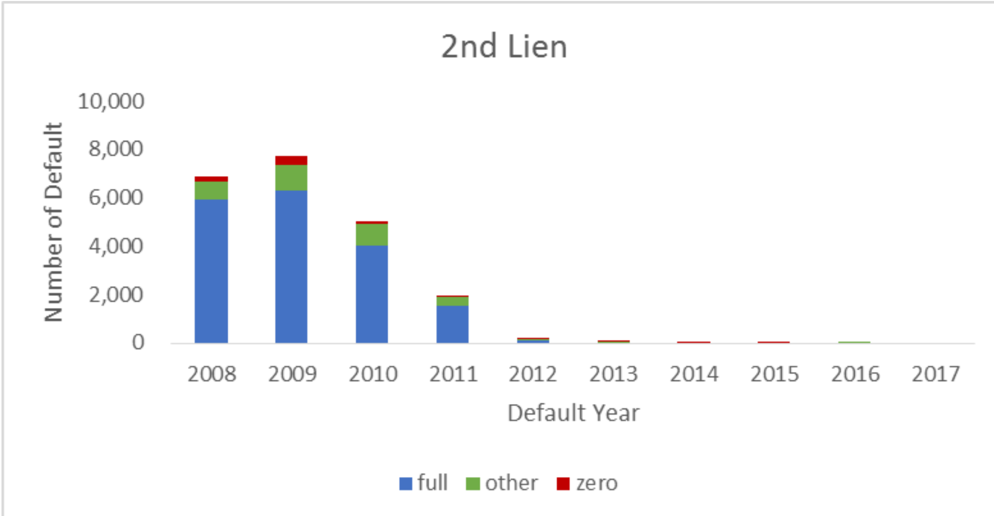
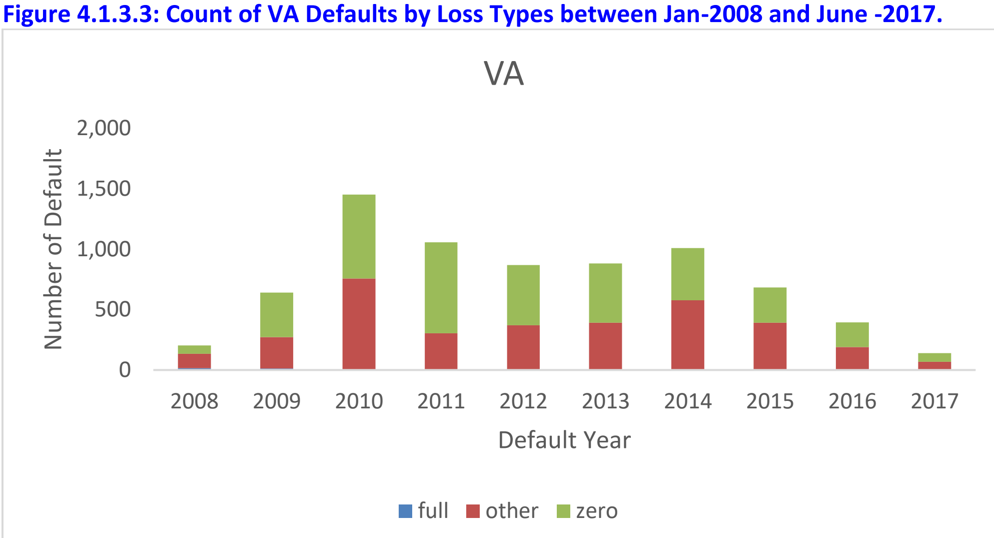


Figure 4.1.3.3: Count of VA Defaults by Loss Types between Jan-2008 and June -2017.



The modeling team provides a data dictionary and the detail of summary statistics for all input variables used in the Method A RM Model suite. Overall, the predictor variables can be divided into three categories: macroeconomic variables, loan-level variables, and bureau variables. A macroeconomic variable, also known as a 'state' variable, changes over time. Account-level loan attributes are either constant over time or time-invariant. Certain time-varying account-level variables, such as current CLTV, are derived by combining a static account-level variable with a dynamic collateral variable. Both decision FICO and refresh FICO are used in the Method A RM model, the former is time-invariant while the latter changes over time. All these variables have in-depth description listed in the data dictionary and the statistical summary tables listed as below. Please refer to the Gating Principles within Chapter 3 for additional details on the listed variables.

To warrant high data quality, CAMU performs extreme value analysis to compliment business insight. When missing or outliers are identified, appropriate treatments are provided case by case. For instance, to smooth outliers, CAMU imputes data with caps and floors to keep the data fell within a reasonable range. After imputation, CAMU will inspect if there is any exclusion bias by examining the extreme values and the deviation from the mean.   If no severe deviation found, it indicates the impact is not significant and the data is reliable. For scores that fall out of the range, they are replaced by the median value and a dummy variable is created to identify the imputation. Detail about the logic, reason of clean up, and the value before and after imputation are described in section 4.1.6. Please refer to attachment '4.1.3 - RM_Method_A_Data_Dictionary.xlsx' for the RM PD model data dictionary

Table 4.1.3.13: Data Dictionary of Method A RM PD Model with Model Drivers

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
ARM_5Yr	arm for 5 yr.	fixed_rate_ind, arm_type_year	B_S_ARM_IND=FIXED_RATE_IND='N'; ARM_5Yr=B_S_ARM_IND=1 and ARM_TYPE_YEAR in ('5YR');
ARM_GT5Yr	arm > 5yr	fixed_rate_ind, arm_type_year	ARM_GT5Yr=B_S_ARM_IND=1 and (ARM_LE3Yr^=1) and (ARM_5Yr^=1);
B_M_coborrower_IND	indicator of coborrower	ssn_id	if in_sec=1 or ssn_id="1" then B_M_coborrower='1'; else B_M_coborrower='0';
B_M_EVER_MBA_60P_IND	indicator of ever mba past 60 days	ever_mba_60p_ind	
B_M_IN_TRIAL_IND	indicator of trial loan	Trial_init_mon, Trial_comp_mon	B_M_IN_TRIAL_IND= (.<TRIAL_INIT_MON_ <= file_dt <= TRIAL_COMP_MON_ or (.<TRIAL_INIT_MON_ <= file_dt and TRIAL_COMP_MON_=.));
B_S_ARM_IND	indicator of arm loans	fixed_rate_ind	renamed arm_ind = B_S_ARM_IND

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
B_S_BROK_IND	indicator of channel via broker	channel_cd	B_S_BROK_IND=C_S_CHANNEL_CD='BROK';
B_S_CORR_IND	indicator of channel via correspondence	channel_cd	B_S_CORR_IND=C_S_CHANNEL_CD='CORR';
B_S_DocFull_IND	indicator of income in full at application	income_doc_cd	if INCOME_DOC_CD in ("CSTD","FULL") then B_S_DocFull='1'; else B_S_DocFull='0';
B_S_DocLow_IND	indicator of low income	income_doc_cd	
B_S_DocMISS_IND	indicator of missing income document at application	income_doc_cd	if INCOME_DOC_CD in (" ",".") then B_S_DocMISS='1'; else B_S_DocMISS='0';

B_S_ge7YR_ARM_IND	indicator of arm year >= 7Yr	arm_type_year, orig_term	B_S_ge7YR_ARM_IND = 0;
			if Fixed_Rate_IND='Y' then do;
			if N_S_ORIG_TERM<=180 then B_S_le15YR_FRM_IND = 1;else B_S_gt15YR_FRM_IND = 1;
			end;
			else do;
			if ARM_TYPE_YEAR in ('1YR','2YR','3YR','0MO','1MO','2MO','3MO','6MO','4YR') then B_S_lt5YR_ARM_IND=1;else
			if ARM_TYPE_YEAR = '5YR' then B_S_5YR_ARM_IND = 1;else
			if ARM_TYPE_YEAR in ('7YR','8YR','OTHER','10YR','6YR','') then B_S_ge7YR_ARM_IND = 1;
			end;

B_S_gt15YR_FRM_IND	indicator or fixed rate term >= 15 Yr	arm_type_year, orig_term	B_S_ge7YR_ARM_IND = 0;
			if Fixed_Rate_IND='Y' then do;
			if N_S_ORIG_TERM<=180 then B_S_le15YR_FRM_IND = 1;else B_S_gt15YR_FRM_IND = 1;
			end;
			else do;
			if ARM_TYPE_YEAR in ('1YR','2YR','3YR','0MO','1MO','2MO','3MO','6MO','4YR') then B_S_lt5YR_ARM_IND=1;else
			if ARM_TYPE_YEAR = '5YR' then B_S_5YR_ARM_IND = 1;else
			if ARM_TYPE_YEAR in ('7YR','8YR','OTHER','10YR','6YR','') then B_S_ge7YR_ARM_IND = 1;
			end;

B_S_INT_ONLY_IND	indicator of trial loan	Trial_init_mon, Trial_comp_mon	B_M_IN_TRIAL_IND= (.<TRIAL_INIT_MON_ <= file_dt <= TRIAL_COMP_MON_ or (.<TRIAL_INIT_MON_ <= file_dt and TRIAL_COMP_MON_=.));

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)

B_S_le15YR_FRM_IND	Indicator of fixed rate term <= 15 yr	arm_type_year, orig_term	if Fixed_Rate_IND='Y' then do;
			if N_S_ORIG_TERM<=180 then B_S_le15YR_FRM_IND = 1;else B_S_gt15YR_FRM_IND = 1;
			end;

B_S_lienfirst_IND	indicator of first lien	lien_cd	Lienfirst_ind=(B_S_lien_CD='0');
B_S_lien2nd_IND	indicator of 2nd lien	lien_cd	Lien2nd_ind=(B_S_lien_CD='1');
B_S_LoanPurpMiss_IND	indicator of missing loan purchase info	loan_purp_cd	if LOAN_PURP_CD in ("."," ") then B_S_LoanPurpMiss='1'; else B_S_LoanPurpMiss='0';
B_S_LoanPurpPurch_IND	indicator of purchase loans	loan_purp_cd	if LOAN_PURP_CD in ("PURCH") then B_S_LoanPurpPurch='1'; else B_S_LoanPurpPurch='0';
B_S_LoanPurpWork_IND	indicator of work out	loan_purp_cd	if LOAN_PURP_CD in ("WORKO") then B_S_LoanPurpWork='1'; else B_S_LoanPurpWork='0'
B_S_lowIncDoc_IND	indicator of low income	apprved_type_cd	if approval_type_cd in(3,10,12,13,14,18,24,25,53,54,57,60,61,64,70,73,75,76,84,87,90,91,92,93,101,102, 103,104,105,106,107,117,118,119,120,121,122,123,126,128,129,130,131,133,137,138,200,204,205,206) or approval_type_cd in(85,86,88,89,124,201,202) then B_S_lowIncDoc='1'; else B_S_lowIncDoc='0';/*11/19/2016 Shuangxi changed*/

B_S_missS_M_FicoRefresh	Indicator of both decision fico and refresh fico are missing	Dcsn_fico, beacon50_score	B_S_missS_M_FicoRefresh = 0;
			S_M_FicoRefresh2 =S_M_FicoRefresh;
			if miss_S_S_FicoDcsn = 1 and miss_S_M_FicoRefresh = 1 then B_S_missS_M_FicoRefresh = 1;
			else if miss_S_S_FicoDcsn = 0 and miss_S_M_FicoRefresh = 1 then do;
			if N_M_MOB <= 3 then S_M_FicoRefresh2 = S_S_FicoDcsn;
			else do;
			S_M_FicoRefresh2 = S_S_FicoDcsn;
			B_S_missS_M_FicoRefresh = 1;
			end;
			end;

B_S_OCC_IV_IND	indicator of occupancy = investment	occ_type_cd	B_S_OCC_IV_IND=C_S_OCC_TYPE_CD='IV';
B_S_OCC_OO_IND	indicator of occupancy = owner occupied	occ_type_cd	B_S_OCC_OO_IND=C_S_OCC_TYPE_CD='OO';
B_S_OCC_SH_IND	indicator of occupancy = second home	occ_type_cd	B_S_OCC_SH_IND=C_S_OCC_TYPE_CD='SH';
B_S_pmi_IND	indicator of private mortage insurance coverage	loan_type_cd	if loan_type_cd=8 then B_S_pmi='1'; else B_S_pmi='0';

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
B_S_PropCondoCP_IND	indicator of property = condo	prop_type_cd	if PROP_TYPE_CD in ("COOP","COND","MAFG") then B_S_PropCondoCP='1'; else B_S_PropCondoCP='0';
B_S_PropMulti_IND	indicator of property = multiple	prop_type_cd	if PROP_TYPE_CD in ("2FAM","3FAM","4FAM","5FAM","MULT") then B_S_PropMulti='1'; else B_S_PropMulti='0';
B_S_PropSFR_IND	indicator of single family	prop_type_cd	if PROP_TYPE_CD in ("SFR") then B_S_PropSFR='1'; else B_S_PropSFR='0';
B_S_WLST_IND	indicator of channel via WLST	channel_cd	B_S_WLST_IND=C_S_CHANNEL_CD='WLST';
B_S_WLST_IND_2nd			B_S_WLST_IND_2nd=B_S_WLST_IND*B_S_lien2nd_IND;
Balloon	indicator of balloon loan	balloon_ind	Balloon=(balloon_ind='Y');
Balloon_Mature	Indicator of balloon loan after loan is mature	balloon_ind, N_M_MOB, N_S_Orig_term	Balloon_Mature=Balloon*(N_M_MOB_cap>=N_S_ORIG_TERM);
community_loan_ind	Indicator or community loans		
cpb_ind	indicator CPB loan		
curr_neg	HPI in negative	HPI	
D_M_PRIN_BAL_10k_SP5LSP15_CMI	principle balance for CPB loans capped at 150K and floored at 50K	Prin_bal	D_M_PRIN_BAL_in10k_SP5LSP15 = min(max(D_M_PRIN_BAL_in10k,5),15); D_M_PRIN_BAL_10k_SP5LSP15_CMI = D_M_PRIN_BAL_in10k_SP5LSP15*(cpb_ind=0);
D_M_PRIN_BAL_in10K	principle balance in 10K	Prin_bal	D_M_PRIN_BAL_in10K=D_M_PRIN_BAL/1e4;
D_M_PRIN_BAL_in10K_CMI_first	principle balance in 10K for first lien loans (non-CPB)	Prin_bal	D_M_PRIN_BAL_in10K_CMI_first =D_M_PRIN_BAL_in10K_CMI*(B_S_Lien2nd_IND=0);
D_M_PRIN_BAL_in10K_CMI_2nd	principle balance in 10K for 2nd lien loans (non-CPB)	Prin_bal	D_M_PRIN_BAL_in10K_CMI_2nd =D_M_PRIN_BAL_in10K_CMI*(B_S_Lien2nd_IND=1);
D_M_PRIN_BAL_in10K_CMI_CONV	principle balance in 10K for non GOV (non-CPB)	Prin_bal	D_M_PRIN_BAL_in10K_CMI=D_M_PRIN_BAL_in10K*(CPB_IND=0); D_M_PRIN_BAL_in10K_CMI_CONV=D_M_PRIN_BAL_in10K_CMI*(GOV=0);
D_M_PRIN_BAL_in10K_CMI_HPIDecr	principle balance in 10K for non CPB loans with HPI in past 12mos <= 1 (non-CPB)	Prin_bal	D_M_PRIN_BAL_in10K_CMI_HPIDecr =D_M_PRIN_BAL_in10K_CMI*(HPI_tm12_ratio<=1);
D_M_PRIN_BAL_in10k_CPB	principle balance in 10k for CPB loans	Prin_bal	D_M_PRIN_BAL_in10K_CPB=D_M_PRIN_BAL_in10K*CPB_IND;
D_M_PRIN_BAL_in10k_LSP10	spline principle balance in 10k <= 100K	Prin_bal	D_M_PRIN_BAL_in10K_LSP10=min(D_M_PRIN_BAL_in10K,10)
D_M_PRIN_BAL_in10k_LSP20_CMICONV	spline principle balance in 10k <= 200K (non-CPB and non-GOV)	Prin_bal	D_M_PRIN_BAL_in10k_LSP20 = min(D_M_PRIN_BAL_in10k,20); D_M_PRIN_BAL_in10k_LSP20_CMI = D_M_PRIN_BAL_in10k_LSP20*(cpb_i

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
			nd=0); D_M_PRIN_BAL_in10k_LSP20_CMICONV = D_M_PRIN_BAL_in10k_LSP20_CMI*(GOV=0);
D_M_PRIN_BAL_in10k_SP10_GOV	spline principle balance in 10K >=100K (GOV)	Prin_bal, portfolio_type_cd	D_M_PRIN_BAL_in10K_SP10=max(D_M_PRIN_BAL_in10K,10); D_M_PRIN_BAL_in10K_SP10_GOV = D_M_PRIN_BAL_in10K_SP10*(GOV=0);
D_M_PRIN_BAL_in10k_SP20_CMICONV	spline principle balance in >=200K (non-CPB and non-GOV)	Prin_bal, portfolio_type_cd	D_M_PRIN_BAL_in10k_SP20_CMICONV = D_M_PRIN_BAL_in10k_SP20_CMI**(GOV=0);
D_M_PRIN_BAL_LE2K	principle balance <= 2K	Prin_bal	D_M_PRIN_BAL_LE2K=D_M_PRIN_BAL_in10K<=0.2;
D_M_PRIN_BAL_LE5K	principle balance <= 5K	Prin_bal	D_M_PRIN_BAL_LE5K=D_M_PRIN_BAL_in10K<=0.5;
D_M_PRIN_BAL_SP100K_CMI_first	principle balance >=100K (non-CPB and non-GOV)	Prin_bal	D_M_PRIN_BAL_SP100K_CMI = D_M_PRIN_BAL_SP100K*(cpb_ind=0); D_M_PRIN_BAL_SP100K_CMI_first =D_M_PRIN_BAL_SP100K_CMI*(B_S_Lien2nd_IND=0);
D_M_PRIN_BAL_SP100K_CMI_HPIDecr	spline principle balance in 10K > 100K for non CPB loans with HPI in past 12mos <= 1 (non-CPB)	Prin_bal , lien_cd	D_M_PRIN_BAL_SP100K_CMI_HPIDecr =D_M_PRIN_BAL_SP100K_CMI*(HPI_tm12_ratio<=1);
DEC	Indicator of month Dec	file_dt	MON=MONTH(FILE_DT); DEC=MON=12;
FCL_Moratorium_Dec09	foreclosure moratorium is enacted in Dec'09	file_dt	FCL_Moratorium_JF09 =('01jan2009'd<=file_dt<='01feb2009'd);
FCL_Moratorium_JF09	foreclosure moratorium is enacted between Jan'09 and Feb'09	file_dt	FCL_Moratorium_Dec09 = file_dt = mdy(12,1,2009);
fcl_time	foreclosre time		
fcl_time_CONV	foreclosre time interacts with non-GOV indicator	fcl_sale_dt, fcl_begin_dt, portfolio_type_cd	fcl_time_CONV = fcl_time*(GOV=0);
FEB	indicator of month Februrary	file_dt	MON=MONTH(FILE_DT); FEB=MON=2;
FRM_LE15Yr	fixed rate term <=15 Yr	fixed_rate_ind	FRM_LE15Yr=B_S_ARM_IND=0 and N_S_ORIG_TERM<=180;
GOV	indicator of GOV loans	portfolio_type_cd	GOV=C_M_PORTFOLIO_TYPE_CD='GOV';
HPI_APP	HPI index (app) associates with time of application	HPI	HPI_APP=HPI_M0_v4/HPI_Orig_v4;
HPI_APP_LE1	HPI index (app) <=1	HPI	HPI_APP_LE1=HPI_APP<=1;
HPI_APP_LSP1_CMI	HPI index (app) <=1 for non-CPB loans	HPI	HPI_APP_LSP1_CMI =HPI_APP_LSP1*(CPB_IND=0);
HPI_tm12_ratio	HPI index (tm12) associates with past 12 mos	HPI	
HPI_tm12_ratio_CMI	HPI index (tm12) associates with past 12 mos for non-CPB	HPI	HPI_tm12_ratio_CMI= HPI_tm12_ratio*(CPB_IND=0);

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
HPI_tm12_ratio_LE1	indicator of HPI index (tm12) <=1	HPI	HPI_TM12_RATIO_LE1=HPI_TM12_RATIO<=1;
HPI_tm12_ratio_LSP1	HPI index (tm12) <=1	HPI	HPI_tm12_ratio_LSP1=min(HPI_tm12_ratio,1);
HPI_tm12_ratio_LSP1_CMI_CONV	HPI index (tm12) <=1 interacts with non-CPB and non-GOV loans indicators	HPI, portfolio_type_cd	HPI_tm12_ratio_LSP1_CMI= HPI_tm12_ratio_LSP1*(CPB_IND=0); HPI_tm12_ratio_LSP1_CMI_CONV = HPI_tm12_ratio_LSP1_CMI*(GOV=0);
HPI_tm12_ratio_SP1	HPI index (tm12) >=1	HPI	HPI_tm12_ratio_SP1=max(HPI_tm12_ratio,1);
HPI_tm12_ratio_SP1_2nd	HPI index (tm12) >=1 for 2nd lien	HPI, lien_cd	HPI_tm12_ratio_SP1_2nd = HPI_tm12_ratio_SP1*(B_S_Lien2nd_IND=1);
HPI_tm12_ratio_SP1_MOBGT120	interaction of HPI index (tm12) with MOB > 120	HPI	HPI_tm12_ratio_SP1_MOBGT120=HPI_tm12_ratio_SP1*N_M_MOB_GT120;
HPI_tm12_ratio_WLST_CMI	HPI index (tm12) assocates with non-CPB loans via Wall Street channel	HPI, channel_cd	B_S_WLST_IND=C_S_CHANNEL_CD='WLST'; HPI_tm12_ratio_WLST=HPI_tm12_ratio*B_S_WLST_IND; HPI_tm12_ratio_WLST_CMI = HPI_tm12_ratio_WLST*(CPB_IND=0);
income_12m_ratio	Ratio of Income in past 12 mos	income	if income_12m^=. then income_12m_ratio=income/income_12m;
income_12m_ratio_LSP1	Ratio of income in past 12 mos with spine <= 1	income	income_12m_ratio_LSP1=min(income_12m_ratio,1);
INX_SP500_12m_ratio	ratio of SP500 in past 12 mos	SP500	if INX_SP500_12m^=. then INX_SP500_12m_ratio=INX_SP500/INX_SP500_12m;
INX_SP500_12m_ratio_LSP1	ratio of SP500 in past 12 mos with spine <= 1	SP500	INX_SP500_12m_ratio_LSP1 = min(INX_SP500_12m_ratio,1);
INX_SP500_12m_ratio_SP1		SP500	INX_SP500_12m_ratio_SP1=max(INX_SP500_12m_ratio,1);
INX_SP500_3m_ratio	ratio of SP500 in past 3 mos	SP500	INX_SP500_3m=lag3(INX_SP500); if INX_SP500_3m^=. then INX_SP500_3m_ratio=INX_SP500/INX_SP500_3m;
INX_SP500_3m_ratio_CONV	ratio of SP500 in past 3 mosfor non-GOV loans	SP500, portfolio_type_cd	INX_SP500_3m_ratio_CONV= INX_SP500_3m_ratio*(GOV=0);
INX_SP500_tm_CONV	SP500 of the month for non-GOV loans	SP500, portfolio_type_cd	INX_SP500_tm =INX_SP500; INX_SP500_tm_CONV=INX_SP500_tm*(GOV=0);
JAN	indicator of month January	file_dt	MON=MONTH(FILE_DT); JAN=MON=1;
judicial	judicial states	prop_state_cd	judicial= (C_S_PROP_STATE_CD2 in ('CT', 'DE', 'FL', 'IL', 'IN', 'IA', 'KS', 'KY', 'LA', 'MA', 'MD', 'ME', 'ND', 'NJ', 'NM', 'NY', 'OH', 'OK', 'PA', 'PR', 'SC', 'SD', 'UT', 'VT', 'VI', 'WI'));
judicial_CONV	Indicator of judicial state associates with non-GOV loans	prop_state_cd, portfolio_type_cd	judicial_CONV = judicial*(GOV=0);

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
JUL	indicator of month July	file_dt	MON=MONTH(FILE_DT); JUL=MON=7;
JUN	indicator of month June	file_dt	JUN=MON=6;
Katrina	indicator of area covered by storm Katrina	prop_zip_cd	Katrina=(Katrina_area and mdy(8,1,2005)<=file_dt<=mdy(7,1,2007));
MAR	Indicator of month March	file_dt	MAR=MON=3;
MAY	indicator of month May	file_dt	MAY=MON=5;
miss_N_M_Age_Oldest_Mtg	indicator of missing related info	cmi16	miss_N_M_Age_Oldest_Mtg = missing(N_M_Age_Oldest_Mtg);
miss_N_M_age_oldest_trd	indicator of missing related info	var_k	miss_N_M_age_oldest_trd=missing(N_M_age_oldest_trd);
miss_N_M_n_open_mort	indicator of missing related info	var_a39	miss_N_M_n_open_mort=missing(N_M_n_open_mort);

MTM_CLTV	ratio of MTM to CLTV	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_LTV=D_M_PRIN_BAL/(D_S_ORIG_PROP_AMT*HPI_M0/HPI_ORIG)*100;
			MTM_CLTV=MTM_LTV;
			if B_S_lienfirst^='1' then do;
			if D_S_OrigBalance>0 and D_S_FST_MTG_BAL>0 then do;
			
			Junior_Ratio_Orig=min(max(0,D_S_OrigBalance/(D_S_FST_MTG_BAL+1e-6)),10);
			
			MTM_CLTV=MTM_LTV*(1+1/Junior_Ratio_Orig);
			end;else
			
			MTM_CLTV=MTM_LTV*(P_S_OrigCLTV/P_S_OrigLTV);
			end;

MTM_CLTV_GT100_GOV	Market to market CLTV with spline > 100 (GOV)	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV, portfolio_type_cd	MTM_CLTV_GT100_GOV=(MTM_CLTV>100)*(GOV=1);
MTM_CLTV_LE10	indicator of arket to market CLTV <=10	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_LE10=MTM_CLTV<=10;
MTM_CLTV_LSP100	Market to market CLTV with spline <=100	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_LSP100=min(MTM_CLTV,100);
MTM_CLTV_LSP60	Market to market CLTV with spline <=60	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_LSP60=min(MTM_CLTV,60);
MTM_CLTV_MOBGT120	interaction of market to market CLTV with MOB >120	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV,	MTM_CLTV_MOBGT120=MTM_CLTV*N_M_MOB_GT120;

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
		origLTV	
MTM_CLTV_SP40_first_CONV	Market to market CLTV with spline >=40 for non-CPB first lien loans	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV, portfolio_type_cd	MTM_CLTV_SP40=max(MTM_CLTV,40); MTM_CLTV_SP40_first = MTM_CLTV_SP40 * B_S_lienfirst_IND; MTM_CLTV_SP40_first_CONV= MTM_CLTV_SP40_first*(GOV=0);
MTM_CLTV_SP40LSP120	Market to market CLTV with spline capped at 120 and floored at 40	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_SP40LSP120=min(max(MTM_CLTV,40),120);
MTM_CLTV_SP40LSP120_GOV	Market to market CLTV with spline capped at 120 and floored at 40 (GOV)	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV, portfolio_type_cd	MTM_CLTV_SP40LSP120_GOV=MTM_CLTV_SP40LSP120*GOV
MTM_CLTV_SP60	Market to market CLTV with spline >=60	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_SP60=max(MTM_CLTV,60);
MTM_CLTV_SP60_HPIDcr	Market to market CLTV >=60 with HPI in past 12mos <= 1	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_SP60_HPIDcr = MTM_CLTV_SP60*(HPI_tm12_ratio<=1);
MTM_CLTV_SP60LSP80_CPB	Market to market CLTV capped at 80 and floored at 60 for CPB loans	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_SP60LSP80_CPB = MTM_CLTV_SP60LSP80*cpb_ind;
MTM_CLTV_SP80	Market to market CLTV with spline >=80	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	MTM_CLTV_SP80=max(MTM_CLTV,80);
MTM_CLTV_SP80_first_CONV	Market to market CLTV with spline >=80 for non-GOV first lien loan	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV, portfolio_type_cd	MTM_CLTV_SP80_first = MTM_CLTV_SP80 * B_S_lienfirst_IND MTM_CLTV_SP80_first_CONV= MTM_CLTV_SP80_first*(GOV=0);
MTM_CLTV_SP80_2nd_CONV	Market to market CLTV with spline >=80 for non-GOV 2nd lien loan	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV, portfolio_type_cd	MTM_CLTV_SP80_2nd_CONV= MTM_CLTV_SP80_2nd*(GOV=0);
N_M_Age_Oldest_Mtg	age of oldest mortgage (bureau)	cmi16	N_M_Age_Oldest_Mtg = input(CMI16,20.);
N_M_age_oldest_trd	age of oldest trade (bureau)	var_k	age_oldest_trd=input(var_k,20.);
N_M_Curr_lt_Orig_Rate_LSP48	current note rate > origination rate with spline <= 48	curr_note_rate,	N_M_Curr_lt_Orig_Rate_LSP48 = min(N_M_Curr_lt_Orig_Rate,48);
N_M_logHpiChange_Orig	change of HPI at origination with log transformation	HPI	if (hpi_m0*hpi_orig > 0) then N_M_logHpiChange_Orig=log(hpi_m0/hpi_orig); else N_M_logHpiChange_Orig = 1;

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
N_M_logHpiChange_Orig_LSP0	change of HPI associates with the origination with log transformation. Spline sets to <= 0	HPI	N_M_logHpiChange_Orig_LSP0=min(N_M_logHpiChange_Orig,0);
N_M_logHpiChange_Orig_SP0	change of HPI associates with the origination with log transformation. Spline sets to >= 0	HPI	N_M_logHpiChange_Orig_SP0=max(N_M_logHpiChange_Orig,0);
N_M_MOB_cap_WLST_CONV	MOB with cap associates with non-GOV loans from WLST channel	loan_orig_dt, file_dt, portfolio_type_cd	N_M_MOB_cap=min(N_M_MOB,180); N_M_MOB_cap_WLST=N_M_MOB_cap*B_S_WLST_IND; N_M_MOB_cap_WLST_CONV= N_M_MOB_cap_WLST*(GOV=0);
N_M_MOB_LSP36	MOB <= 36	loan_orig_dt, file_dt	N_M_MOB_LSP36=min(N_M_MOB_cap,36);
N_M_MOB_LSP36_GOV	MOB <= 36 (GOV)	loan_orig_dt, file_dt, portfolio_type_cd	N_M_MOB_LSP36_GOV=N_M_MOB_LSP36*GOV;
N_M_MOB_LSP72	MOB <= 72	loan_orig_dt, file_dt	N_M_MOB_LSP72=min(N_M_MOB_cap,72);
N_M_MOB_SP36	MOB >= 36	loan_orig_dt, file_dt	N_M_MOB_SP36=max(N_M_MOB_cap,36);
N_M_MOB_SP36LSP180	MOB capped at 180 and floored at 36.	loan_orig_dt, file_dt	N_M_MOB_SP36LSP180=min(max(N_M_MOB,36),180);
N_M_n_open_mort	number of open mortgage (bureau)	var_a39	n_open_mort=input(var_a39,20.);
never_neg	indicator for HPI before negative	hpi	never_neg=((.<file_dt<firstdate_neg) or firstdate_neg=.);
OCT	indicator of month October		OCT=MON=10;
P_AND_I_ratio_ARM	ratio of principal interest amount with the amount in past 6 mos for ARM loans	principal_interest_amt, curr_note_rate, prin_bal, fixed_rate_ind	P_AND_I_ratio_ARM=P_AND_I_ratio*B_S_ARM_IND;
P_AND_I_ratio_IO	ratio of principal interest amount with the amount in past 6 mos for loans with IO	principal_interest_amt, curr_note_rate, prin_bal, int_only_ind	P_AND_I_ratio_IO=P_AND_I_ratio*B_S_INT_ONLY_IND;
P_M_CURR_NOTE_RATE_ARM	Interaction of current note rate and ARM indicator	curr_note_rate, fixed_rate_ind	P_M_CURR_NOTE_RATE_ARM=P_M_CURR_NOTE_RATE*B_S_ARM_IND;
P_M_CURR_NOTE_RATE_ARM_CONV	Interaction of current note rate and ARM indicator for non-GOV loans	curr_note_rate, fixed_rate_ind, portfolio_type_cd	P_M_CURR_NOTE_RATE_ARM=P_M_CURR_NOTE_RATE*B_S_ARM_IND; P_M_CURR_NOTE_RATE_ARM_CONV =P_M_CURR_NOTE_RATE_ARM*(GOV=0);
P_M_PresIntSpread_LSP3pt5_CMI	Interest spread <=3.5 for non-CPB	curr_note_rate, swap	P_M_PresIntSpread_LSP3pt5 = min(P_M_PresIntSpread,3.5); P_M_PresIntSpread_LSP3pt5_CMI = P_M_PresIntSpread_LSP3pt5*(cpb_ind =0);
P_M_PresIntSpread_LSP3pt5_FRM	Interaction of Interest spread with fixed rate indicator	curr_note_rate, swap, FIXED_RATE_IND	P_M_PresIntSpread_LSP3pt5_FRM = P_M_PresIntSpread_LSP3pt5*B_S_FIXED_RATE_IND;
P_M_PresIntSpread_LSP5_CPB	Interest spread <=5 (CPB)		P_M_PresIntSpread_LSP5 = min(P_M_PresIntSpread,5); P_M_PresIntSpread_LSP5_CPB = P_M_PresIntSpread_LSP5*cpb_ind;

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
P_M_PresIntSpread_LSP5_PRA	interest spread spline interacts with pre-arm reset indicator		P_M_PresIntSpread_LSP5_PRA = P_M_PresIntSpread_LSP5*(B_S_PreResetArm_IND);
P_M_PresIntSpread_LSP5_Pre2010	interest spread spline interacts with prior 2010 indicator		P_M_PresIntSpread_LSP5_Pre2010 = P_M_PresIntSpread_LSP5*(Post2010_Orig=0);
P_M_State_Unemp_Rate	unemployment rate in state level	unemployment	
P_M_State_Unemp_Rate_2nd	Unemployment in State level interacts with 2nd lien indicator	unemployment, lien_cd	P_M_State_Unemp_Rate_2nd=P_M_State_Unemp_Rate*(B_S_Lien2nd_IND=1);
P_M_State_Unemp_Rate_currneg	Unemployment in State level interacts with current non-negative indicator	unemployment, HPI	P_M_State_Unemp_Rate_currneg=P_M_State_Unemp_Rate*curr_neg.
P_M_State_Unemp_Rate_lag12_GOV	unemployment in State level with 12 mos lag interacts with GOV indicator	unemployment, portfolio_type_cd	P_M_State_Unemp_Rate_lag12=P_M_State_Unemp_Rate/R_M_State_UnempB12M; P_M_State_Unemp_Rate_lag12_GOV = P_M_State_Unemp_Rate_lag12*GOV;
P_S_BE_DEBT_RATIO	ratio of income paying debt	be_debt_ratio	rename BE_DEBT_RATIO = P_S_BE_DEBT_RATIO;
P_S_OrigIntSpread	interest spread at origination	orig_note_rate, swap	P_S_OrigIntSpread=orig_note_rate - market_O;
P_S_OrigIntSpread_FRM	interest spread interacts with fixed rate indicator	orig_note_rate, swap, fixed_rate_ind	P_S_OrigIntSpread_FRM=P_S_OrigIntSpread*(1-B_S_ARM_IND);;
POST_ARM_12M_IND	indicator of 6 mos before ARM	arm_type_year	if B_S_ARM_IND = 1 and POST_ARM_IND = 1 and -12<=ARM_MonthLeft<=0 then POST_ARM_12M_IND = 1;
POST_ARM_6M_IND	indicator of 6 mos before ARM	arm_type_year	IF ARM_TYPE_YEAR = '1YR' THEN RESET_period=12; ELSE IF ARM_TYPE_YEAR = '1MO' THEN RESET_period=1; ELSE IF ARM_TYPE_YEAR = '10YR' THEN RESET_period=120; ELSE IF ARM_TYPE_YEAR = '2YR' THEN RESET_period=24; ELSE IF ARM_TYPE_YEAR = '2MO' THEN RESET_period=2; ELSE IF ARM_TYPE_YEAR = '3YR' THEN RESET_period=36; ELSE IF ARM_TYPE_YEAR = '3MO' THEN RESET_period=3; ELSE IF ARM_TYPE_YEAR = '4YR' THEN RESET_period=48; ELSE IF ARM_TYPE_YEAR = '5YR' THEN RESET_period=60; ELSE IF ARM_TYPE_YEAR = '6YR' THEN RESET_period=72; ELSE

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)

			IF ARM_TYPE_YEAR = '6MO' THEN RESET_period= 6; ELSE
			IF ARM_TYPE_YEAR = '7YR' THEN RESET_period=84; ELSE
			IF ARM_TYPE_YEAR = '8YR' THEN RESET_period=96; ELSE
			IF ARM_TYPE_YEAR ='OTHER' THEN RESET_period= 1; ELSE
			
			RESET_period=999;
			if RESET_period=999 then ARM_MonthLeft=999;
			else ARM_MonthLeft=RESET_period-N_M_MOB;
			
			POST_ARM_IND=.<ARM_MonthLeft<=0;
			POST_ARM_6M_IND=-6<ARM_MonthLeft<=0;

POST_ARM_IND	Post arm indicator	arm_type_year	POST_ARM_IND=.<ARM_MonthLeft<=0;
POST_ARM_IND_WLST	interaction of post arm indicator with WLST indicator	fixed_rate_ind, channel_cd	POST_ARM_IND_WLST=POST_ARM_IND*B_S_WLST_IND;
POST_IO_6M_IND	indicator of post IO in 6 mos	int_only_year	IF B_S_INT_ONLY_IND = 1 AND -6<=IO_MonthLeft<=0 then POST_IO_6M_IND = 1;
POST_IO_IND	indicator of post IO period	int_only_year	POST_IO_IND=.<IO_MonthLeft<=0;

post_neg_CONV	HPI change to negative	HPI, portfolio_type_cd	post_neg=0;
			if file_dt>=firstdate_back_pos>. then do;
			post_neg=1;
			end;
			post_neg_CONV=post_neg*(GOV=0);

Post2010_Orig	Origination date after year 2010	loan_orig_dt	Post2010_Orig=N_S_LOAN_Orig_Dt>=mdy(1,1,2010);
Post2010_Orig_CONV	Origination date after year 2010 interacts with non-GOV indicator	loan_orig_dt, portfolio_type_cd	Post2010_Orig_CONV= Post2010_Orig*(GOV=0);
PPP_IND	PPP indicator	PPP_ind	
PRE_ARM_12M_IND	indicator of pre arm in past 12 mos	fixed_rate_ind	PRE_ARM_12M_IND = 0; if B_S_ARM_IND = 1 and POST_ARM_IND = 0 and 0<=ARM_MonthLeft<=12 then PRE_ARM_12M_IND = 1;
PRE_ARM_6M_IND	indicator of pre arm in past 6 mos	fixed_rate_ind	PRE_ARM_6M_IND = 0; if B_S_ARM_IND = 1 and POST_ARM_IND = 0 and 0<=ARM_MonthLeft<=6 then PRE_ARM_6M_IND = 1;

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
PRE_IO_12M_IND	indicator of pre io in past 12 mos	INT_ONLY_IND	PRE_IO_12M_IND = 0; IF B_S_INT_ONLY_IND = 1 AND 0<=IO_MonthLeft<=12 then PRE_IO_12M_IND = 1;
PRE_IO_6M_IND	indicator of pre io in past 6 mos	INT_ONLY_IND	PRE_IO_6M_IND = 0; IF B_S_INT_ONLY_IND = 1 AND 0<=IO_MonthLeft<=6 then PRE_IO_6M_IND = 1;
Pre_Maturity_12M_IND	indicator of pre maturity in past 12 mos	orig_term	Pre_Maturity_12M_IND = 0<= N_S_ORIG_TERM-N_M_MOB<=12;
Pre_Maturity_3M_IND	indicator of pre maturity in past 3 mos	orig_term	Pre_Maturity_3M_IND = 0<= N_S_ORIG_TERM-N_M_MOB<=3;
Pre_Maturity_6M_IND	indicator of pre maturity in past 6 mos	orig_term	Pre_Maturity_6M_IND = 0<= N_S_ORIG_TERM-N_M_MOB<=6;
R_M_fcl_time_B12M_LE1_GOV	interaction of foreclosure past 12 mos indicator with GOV indicator	unemployment, portfolio_type_cd	R_M_fcl_time_B12M_LE1 = R_M_fcl_time_B12M<=1; R_M_fcl_time_B12M_LE1_GOV = R_M_fcl_time_B12M_LE1*GOV;
R_M_State_UnempB12M	unemployment in State level in past 12 mos	Unemployment	if (unemployment_t>0 and unemployment_tm12>0) then R_M_State_UnempB12M=unemployment_t/unemployment_tm12; else R_M_State_UnempB12M = 1;
R_M_State_UnempB12M_BROK_CONV	convicator of r m_state_unempb12m co past BROK mos	unemployment, channel_cd, portfolio_type_cd	R_M_State_UnempB12M_BROK =R_M_State_UnempB12M*B_S_BROK_IND; R_M_State_UnempB12M_BROK_CONV=R_M_State_UnempB12M_BROK*(GOV=0);
R_M_State_UnempB12M_CMI	cmiicator of r m_state cm past UnempB12 mos	unemployment	R_M_State_UnempB12M_CMI =R_M_State_UnempB12M*(CPB_IND =0);
R_M_State_UnempB12M_CORR	corricator of r m_state co past UnempB12 mos	unemployment, channel_cd	R_M_State_UnempB12M_CORR=R_M_State_UnempB12M*B_S_CORR_IND;
R_M_State_UnempB12M_IO_CONV	convicator of r m_state_unempb12m co past IO mos	unemployment, portfolio_type_cd, int_only_year, int_only_ind	R_M_State_UnempB12M_IO =R_M_State_UnempB12M*B_S_INT_ONLY_IND; R_M_State_UnempB12M_IO_CONV= R_M_State_UnempB12M_IO*(GOV=0);
R_M_State_UnempB12M_SP1	unemployment in State level in past 12 mos	unemployment	R_M_State_UnempB12M_SP1 = max(R_M_State_UnempB12M,1);
R_M_State_UnempB12M_SP1_MOBGT120	mobgt120icator of r m_state_unempb12m mo past SP1_MOBGT120 mos	unemployment, loan_orig_dt	R_M_State_UnempB12M_SP1_MOBGT120=R_M_State_UnempB12M_SP1*N_M_MOB_GT120;
R_M_US_GDPB1M	GDP in US level	GDP	

Output variable name	Description and definition	Related fields (if any)	Source field (or formula if calculated field)
S_M_FicoRefresh	refresh fico	beacon50_score	S_M_Fico_P=input(beacon50_score,20.);/*11/21/2016 Shuangxi changed beacon50_score+0*/ S_M_Fico_S=input(beacon50_score_s,20.);/*11/21/2016 Shuangxi changed*/ if in_prim eq 1 and S_M_Fico_P<=0 then B_M_FicoMissing1='1'; else B_M_FicoMissing1='0'; /*11/21/2016 Shuangxi changed*/ if in_prim ne 1 then B_M_FicoMissing2='1'; else B_M_FicoMissing2='0'; /*11/21/2016 Shuangxi changed*/
S_M_FicoRefresh_LSP640	refresh fico floored at 640	beacon50_score	S_M_FicoRefresh_LSP640=min(S_M_FicoRefresh,640);
S_M_FicoRefresh_SP640	refresh fico capped at 640	beacon50_score	S_M_FicoRefresh_SP640=max(S_M_FicoRefresh,640);
S_M_FicoRefresh_SP720	refresh fico capped at 720	beacon50_score	S_M_FicoRefresh_SP720=max(S_M_FicoRefresh,720);
S_S_FicoDcsn	decision fico	Dcsn_fico	S_S_FicoDcsn=0; B_S_FicoDcsnMiss='0';/*11/19/2016 Shuangxi changed 0 to '0' */ IF DCSN_FICO>=300 and DCSN_FICO<=850 then S_S_FicoDcsn=DCSN_FICO;
Sandy	indicator of property in area hit by the storm Sandy	prop_zip_cd	
state_AZ_CONV	Indicator of AZ state	prop_state_cd, portfolio_type_cd	state_AZ = C_S_PROP_STATE_CD2 = 'AZ'; state_AZ_CONV = state_AZ*(GOV=0);
state_FL_CONV	indicator of FL state	prop_state_cd, portfolio_type_cd	state_FL = C_S_PROP_STATE_CD2 = 'FL'; state_FL_CONV = state_FL*(GOV=0);
state_NJ_CONV	indicator of NJ state	prop_state_cd, portfolio_type_cd	state_NJ = C_S_PROP_STATE_CD2 = 'NJ'; state_NJ_CONV = state_NJ*(GOV=0);
state_NY_CONV	indicator of NJ state	prop_state_cd, portfolio_type_cd	state_NY = C_S_PROP_STATE_CD2 = 'NY'; state_NY_CONV = state_NY*(GOV=0);

Table 4.1.3.14: Descriptive Data Summary of PD Model Drivers (Continuous Variable)

Variable	N	N Miss	Minimum	Mean	Maximum	STD
ARM_5Yr	7,557,694	0	0	0.07	1	0.26
ARM_GT5Yr	7,557,694	0	0	0.03	1	0.17
D_M_PRIN_BAL_10k_SP5LSP15_CMI	7,557,694	0	0	10.01	15	4.23
D_M_PRIN_BAL_in10K	7,557,694	0	1E-06	15.15	813.65	20.52
D_M_PRIN_BAL_in10K_CMI_first	7,557,694	0	0	13.89	144.57	15.76
D_M_PRIN_BAL_in10K_CMI_2nd	7,557,694	0	0	0.33	99.48	1.65
D_M_PRIN_BAL_in10K_CMI_CONV	7,557,694	0	0	11.51	144.57	16.27
D_M_PRIN_BAL_in10K_CMI_HPIDecr	7,557,694	0	0	9.55	144.57	14.20
D_M_PRIN_BAL_in10K_CPB	7,557,694	0	0	0.93	813.65	14.35
D_M_PRIN_BAL_in10K_LSP10	7,557,694	0	1E-06	7.65	10	3.05
D_M_PRIN_BAL_in10k_LSP20_CMICONV	7,557,694	0	0	8.22	20	7.42
D_M_PRIN_BAL_in10K_SP10_GOV	7,557,694	0	0	14.44	813.65	20.55
D_M_PRIN_BAL_in10k_SP20_CMICONV	7,557,694	0	0	18.66	144.57	15.29
D_M_PRIN_BAL_SP100K_CMI_first	7,557,694	0	0	15.73	144.57	14.81
D_M_PRIN_BAL_SP100K_CMI_HPIDecr	7,557,694	0	0	10.89	144.57	13.88
fcl_time	7,557,694	0	1	7.79	55	7.59
fcl_time_CONV	7,557,694	0	0	5.76	55	6.85
HPI_APP	7,557,694	0	0.36	1.11	21.13	0.49
HPI_APP_LSP1_CMI	7,557,694	0	0	0.91	1	0.16
HPI_tm12_ratio	7,557,694	0	0.63	0.97	1.43	0.08
HPI_tm12_ratio_CMI	7,557,694	0	0	0.96	1.43	0.12
HPI_tm12_ratio_LSP1	7,557,694	0	0.63	0.95	1	0.06
HPI_tm12_ratio_LSP1_CMI_CONV	7,557,694	0	0	0.73	1	0.40
HPI_tm12_ratio_SP1	7,557,694	0	1	1.02	1.43	0.03
HPI_tm12_ratio_SP1_2nd	7,557,694	0	0	0.08	1.43	0.28
HPI_tm12_ratio_SP1_MOBGT120	7,557,694	0	0	0.16	1.39	0.37
HPI_tm12_ratio_WLST_CMI	7,557,694	0	0	0.22	1.43	0.40
income_12m_ratio	7,557,694	0	0.89	1.01	1.14	0.03
income_12m_ratio_LSP1	7,557,694	0	0.89	0.99	1	0.02
INX_SP500_12m_ratio	7,557,694	0	0.55	1.04	1.50	0.21
INX_SP500_12m_ratio_LSP1	7,557,694	0	0.55	0.93	1	0.13
INX_SP500_12m_ratio_SP1	7,557,694	0	1	1.10	1.50	0.11
INX_SP500_3m_ratio	7,557,694	0	0.70	1.01	1.25	0.10
INX_SP500_3m_ratio_CONV	7,557,694	0	0	0.78	1.25	0.43
INX_SP500_tm_CONV	7,557,694	0	0	1013.25	2673.61	619.89
MTM_CLTV	7,557,694	0	0	76.72	300	36.89

Variable	N	N Miss	Minimum	Mean	Maximum	STD
MTM_CLTV_GT100_GOV	7,557,694	0	0	0.07	1	0.25
MTM_CLTV_LE10	7,557,694	0	0	0.06	1	0.23
MTM_CLTV_LSP100	7,557,694	0	0	71.44	100	28.93
MTM_CLTV_LSP60	7,557,694	0	0	52.08	60	16.19
MTM_CLTV_MOBGT120	7,557,694	0	0	6.36	300	18.57
MTM_CLTV_SP40_first_CONV	7,557,694	0	0	54.26	300	43.76
MTM_CLTV_SP40LSP120	7,557,694	0	40	77.94	120	26.28
MTM_CLTV_SP40LSP120_GOV	7,557,694	0	0	19.05	120	37.36
MTM_CLTV_SP60	7,557,694	0	60	84.63	300	26.63
MTM_CLTV_SP60_HPIDcr	7,557,694	0	0	57.94	300	49.18
MTM_CLTV_SP60LSP80_CPB	7,557,694	0	0	0.62	80	6.37
MTM_CLTV_SP80	7,557,694	0	80	92.46	300	20.81
MTM_CLTV_SP80_first_CONV	7,557,694	0	0	63.44	300	44.80
MTM_CLTV_SP80_2nd_CONV	7,557,694	0	0	7.85	300	27.77
N_M_Age_Oldest_Mtg	7,557,694	0	1	116.82	482	52.85
N_M_age_oldest_trd	7,557,694	0	0	202.39	752	65.95
N_M_Curr_lt_Orig_Rate_LSP48	7,557,694	0	0	3.04	48	10.76
N_M_logHpiChange_Orig	7,557,694	0	-1.01	0.04	3.05	0.33
N_M_logHpiChange_Orig_LSP0	7,557,694	0	-1.01	-0.10	0	0.17
N_M_logHpiChange_Orig_SP0	7,557,694	0	0	0.14	3.05	0.24
N_M_MOB_cap_WLST_CONV	7,557,694	0	0	10.72	180	25.78
N_M_MOB_LSP36	7,557,694	0	1	31.51	36	8.59
N_M_MOB_LSP36_GOV	7,557,694	0	0	7.53	36	14.33
N_M_MOB_LSP72	7,557,694	0	1	50.30	72	22.35
N_M_MOB_SP36	7,557,694	0	36	72.29	180	42.33
N_M_MOB_SP36LSP180	7,557,694	0	36	72.29	180	42.33
N_M_n_open_mort	7,557,694	0	0	1.14	99	0.87
P_AND_I_ratio_ARM	7,557,694	0	0	0.23	2.93	0.42
P_AND_I_ratio_IO	7,557,694	0	0	0.09	2.93	0.28
P_M_CURR_NOTE_RATE_ARM	7,557,694	0	-0.13	1.51	18.75	3.01
P_M_CURR_NOTE_RATE_ARM_CONV	7,557,694	0	-0.13	1.42	18.75	2.97
P_M_PresIntSpread_LSP3pt5_CMI	7,557,694	0	-5.73	2.94	3.5	1.00
P_M_PresIntSpread_LSP3pt5_FRM	7,557,694	0	-5.73	2.33	3.5	1.47
P_M_PresIntSpread_LSP5_CPB	7,557,694	0	-2.56	0.02	5	0.28
P_M_PresIntSpread_LSP5_PRA	7,557,694	0	-5.52	0.46	5	1.35
P_M_PresIntSpread_LSP5_Pre2010	7,557,694	0	-5.73	3.65	5	1.54
P_M_State_Unemp_Rate	7,557,694	0	2.10	7.26	14.6	2.61
P_M_State_Unemp_Rate_2nd	7,557,694	0	0	0.55	14.6	2.00
P_M_State_Unemp_Rate_currneg	7,557,694	0	0	3.91	14.6	4.56
P_M_State_Unemp_Rate_lag12_GOV	7,557,694	0	0	1.61	14.60	3.23

Variable	N	N Miss	Minimum	Mean	Maximum	STD
P_S_BE_DEBT_RATIO	7,557,694	0	0	28.49	100	20.27
P_S_OrigIntSpread	7,557,694	0	-9.41	3.03	12.09	2.10
P_S_OrigIntSpread_FRM	7,557,694	0	-9.41	2.47	12.09	2.30
R_M_State_UnempB12M	7,557,694	0	0.38	1.13	2.53	0.29
R_M_State_UnempB12M_BROK_CONV	7,557,694	0	0	0.10	2.53	0.34
R_M_State_UnempB12M_CMI	7,557,694	0	0	1.12	2.53	0.31
R_M_State_UnempB12M_CORR	7,557,694	0	0	0.44	2.53	0.57
R_M_State_UnempB12M_IO_CONV	7,557,694	0	0	0.10	2.53	0.35
R_M_State_UnempB12M_SP1	7,557,694	0	1	1.18	2.53	0.26
R_M_State_UnempB12M_SP1_MOBGT120	7,557,694	0	0	0.17	2.53	0.41
R_M_US_GDPB1M	7,557,694	0	-0.08	0.03	0.08	0.03
S_M_FicoRefresh	7,557,694	0	300	575.74	850	90.70
S_M_FicoRefresh_LSP640	7,557,694	0	300	559.76	640	64.44
S_M_FicoRefresh_SP640	7,557,694	0	640	655.99	850	38.83
S_M_FicoRefresh_SP720	7,557,694	0	720	724.35	850	15.98
S_S_FicoDcsn	7,557,694	0	300	632.87	850	72.59

Table 4.1.3.15: Descriptive Data Summary of PD Model Drivers (Indicator Variable)

Variable	N	Count (=1)	Count %
B_M_coborrower_IND	7,557,694	3,583,150	47.41%
B_M_IN_TRIAL_IND	7,557,694	218,481	2.89%
B_S_ARM_IND	7,557,694	1,708,771	22.61%
B_S_BROK_IND	7,557,694	820,767	10.86%
B_S_CORR_IND	7,557,694	2,989,494	39.56%
B_S_DocFull_IND	7,557,694	3,648,964	48.28%
B_S_DocLow_IND	7,557,694	722,387	9.56%
B_S_DocMISS_IND	7,557,694	2,107,628	27.89%
B_S_ge7YR_ARM_IND	7,557,694	232,821	3.08%
B_S_gt15YR_FRM_IND	7,557,694	4,884,936	64.64%
B_S_INT_ONLY_IND			8.56%

Variable	N	Count (=1)	Count %
	7,557,694	646,719	
B_S_le15YR_FRM_IND	7,557,694	963,987	12.76%
B_S_lienfirst_IND	7,557,694	6,948,155	91.93%
B_S_lien2nd_IND	7,557,694	609,539	8.07%
B_S_LoanPurpMiss_IND	7,557,694	29,336	0.39%
B_S_LoanPurpPurch_IND	7,557,694	2,737,019	36.22%
B_S_LoanPurpWork_IND	7,557,694	20,646	0.27%
B_S_lowIncDoc_IND	7,557,694	581,680	7.70%
B_S_missS_M_FicoRefresh	7,557,694	254,499	3.37%
B_S_OCC_IV_IND	7,557,694	174,048	2.30%
B_S_OCC_OO_IND	7,557,694	7,293,549	96.50%
B_S_OCC_SH_IND	7,557,694	90,097	1.19%
B_S_pmi_IND	7,557,694	220,638	2.92%
B_S_PropCondoCP_IND	7,557,694	668,513	8.85%
B_S_PropMulti_IND	7,557,694	143,987	1.91%
B_S_PropSFR_IND	7,557,694	6,581,345	87.08%
B_S_WLST_IND	7,557,694	1,748,937	23.14%
B_S_WLST_IND_2nd	7,557,694	249,365	3.30%
Balloon	7,557,694	268,864	3.56%
Balloon_Mature	7,557,694	6,801	0.09%
community_loan_ind	7,557,694	210,243	2.78%
cpb_ind	7,557,694	71,201	0.94%
curr_neg	7,557,694	3,467,751	45.88%
D_M_PRIN_BAL_LE2K	7,557,694	104,228	1.38%
D_M_PRIN_BAL_LE5K	7,557,694	170,220	2.25%

Variable	N	Count (=1)	Count %
DEC	7,557,694	642,979	8.51%
FCL_Moratorium_Dec09	7,557,694	129,920	1.72%
FCL_Moratorium_JF09	7,557,694	209,217	2.77%
FEB	7,557,694	617,761	8.17%
FRM_LE15Yr	7,557,694	963,987	12.76%
GOV	7,557,694	1,675,436	22.17%
HPI_APP_LE1	7,557,694	3,594,934	47.57%
HPI_TM12_RATIO_LE1	7,557,694	4,831,683	63.93%
JAN	7,557,694	562,314	7.44%
judicial	7,557,694	3,738,748	49.47%
judicial_CONV	7,557,694	2,886,385	38.19%
JUL	7,557,694	636,734	8.42%
JUN	7,557,694	639,514	8.46%
MAR	7,557,694	605,059	8.01%
MAY	7,557,694	635,319	8.41%
miss_N_M_Age_Oldest_Mtg	7,557,694	2,076,411	27.47%
miss_N_M_age_oldest_trd	7,557,694	3,553,795	47.02%
miss_N_M_n_open_mort	7,557,694	3,514,185	46.50%
MTM_CLTV_LE10	7,557,694	423,305	5.60%
never_neg	7,557,694	3,823,082	50.59%
OCT	7,557,694	639,846	8.47%
POST_ARM_12M_IND	7,557,694	155,895	2.06%
POST_ARM_6M_IND	7,557,694	79,753	1.06%
POST_ARM_IND	7,557,694	672,837	8.90%
POST_ARM_IND_WLST			2.60%

Variable	N	Count (=1)	Count %
	7,557,694	196,861	
POST_IO_6M_IND	7,557,694	13,133	0.17%
POST_IO_IND	7,557,694	68,869	0.91%
post_neg_CONV	7,557,694	147,996	1.96%
Post2010_Orig	7,557,694	146,710	1.94%
Post2010_Orig_CONV	7,557,694	113,931	1.51%
PPP_IND	7,557,694	926,851	12.26%
PRE_ARM_12M_IND	7,557,694	165,138	2.19%
PRE_ARM_6M_IND	7,557,694	81,995	1.08%
PRE_IO_12M_IND	7,557,694	31,690	0.42%
PRE_IO_6M_IND	7,557,694	16,657	0.22%
Pre_Maturity_12M_IND	7,557,694	79,808	1.06%
Pre_Maturity_3M_IND	7,557,694	31,316	0.41%
Pre_Maturity_6M_IND	7,557,694	47,575	0.63%
R_M_fcl_time_B12M_LE1_GOV	7,557,694	1,011,091	13.38%
state_AZ_CONV	7,557,694	115,187	1.52%
state_FL_CONV	7,557,694	479,720	6.35%
state_NJ_CONV	7,557,694	135,589	1.79%
state_NY_CONV	7,557,694	334,913	4.43%

Please refer to attachment ‘4.1.3 Severity Method A Data Dictionary.xlsx’ for Severity model data dictionary.

Table 4.1.3.16: Descriptive Data Summary of LGD Continuous Model Drivers

Variable	N	N Miss	Minimum	Mean	Maximum	Std Dev
first Lien						
cLTV_MTM_unemp_down	65055	0	0	39.04	300	51.06
log_curr_bal	65055	0	0	11.48	16.02	1.41

Variable	N	N Miss	Minimum	Mean	Maximum	Std Dev
log_curr_bal_SP_gt300k	65055	0	12.61	12.68	16.02	0.21
log_curr_bal_SP_le300k	65055	0	0	11.41	12.61	1.33
log_HPI_t_orig_ratio	65055	0	-1.00	-0.16	2.46	0.32
log_HPI_tm12_ratio_SP_le0	65055	0	-0.46	-0.07	0	0.09
LTV_mtm_SP_gt90	65055	0	90	106.83	300	23.96
prin_reduction	65055	0	0	0	0	0
UnempRate	65055	0	0	8.69	14.60	2.34
UnempRate_SP_gt9	65055	0	9	9.83	14.60	1.21
2nd Lien						
cLTV_MTM_SP_gt80_le120	21902	0	80	108.06	120	15.06
junior_ratio_SP_lep20	21902	0	0.01	0.19	0.2	0.03
log_curr_bal_SP_gt10k_le150k	21902	0	9.21	10.53	11.92	0.70
log_curr_bal_SP_gt150k	21902	0	11.92	11.92	12.94	0.05
log_fst_mtg_bal	21902	0	5.96	11.62	14.46	1.00
log_HPI_tm12_ratio	21902	0	-0.46	-0.12	0.21	0.11
prin_reduction	21902	0	0	0	0	0
R_M_UnempB12M_SP_gt1_le1p2	21902	0	1	1.14	1.2	0.08
UnempRate	21902	0	2.70	8.71	14.6	2.29

Table 4.1.3.17: Descriptive Data Summary of LGD Binary Model Drivers

Variable	N	Count	Count %
first Lien			
BROK_IND	65055	8380	12.9%
FHL_ind	65055	0	0.0%
HLC_ind	65055	0	0.0%
judicial	65055	27173	41.8%
pmi	65055	1982	3.0%
2nd Lien			
CBNA_ind	21902	0	0.0%
deficiency	21902	15981	73.0%
judicial	21902	9757	44.5%
lpi	21902	73	0.3%
REL_ind	21902	21902	100.0%
wasUSRB	21902	0	0.0%

If a vendor model is used, describe any model-specific data requirements.

[Describe potential requirements and limitations, including dependencies on legacy formats (e.g., a limit of 65,535 lines per file). Describe how the data was managed to match the requirements or overcome the limitations.]

Not Applicable. The Method A RM model suite is not a vendor model.

1. MRM Question - As per the model limitations, there was no suggestion/comment about changing the selection of OOT sample in the previous limitations from standard practice. Further, based on confirmation/attestation provided by model sponsor, updated method A is supposed to remediate limitation of both Method A and Method B models, except limitation about alternative modeling approach. One of the limitation (limitation id 16856) in FRHEL method B is following "*No consideration of recent 2 years of data either for model development or model stability analysis*". In this limitation also, there was no suggestion/comment about changing the selection of OOT sample from standard practice.

Hence, follow-up questions from MRM are as follows:

- Sponsor needs to further justify creation of this non-standard OOT sample.
- Based on the above limitation (#19546), sponsor is requested to provide justification for not including pre-2008 data in LGD model development sample.

Answer –

<u>Justification for the selection of the OOT Period</u> - In this version of model development, CAMU holds out an OOT sample solely to comply with MRM's OOT test requirement although CAMU believes that holding out OOT data will potentially result in loss of performance and portfolio change information. Due to this concern, CAMU carefully choose Jan-2012 to March-2014 period as out-of-time validation sample and retaining data between Feb 2006 - Dec 2011 and Apr 2014 -Dec 2017 as model development data due to the following considerations –

1. First and foremost, as per the limitation on the prior model suite around the non-inclusion of recent period data, it was recommended to include the most recent time period as part of the model development sample. To comply with MRM's requirements, CAMU did not choose to hold out most recent data as OOT. CAMU also agrees with MRM that holding out recent performance data will disable us from capturing the most recent macro-economic trend, such as interest rate increase, portfolio and underwriting policy change. The inclusion of the recent period data helped in capturing the go-forward state of the business in terms of origination profile and portfolio composition mix.

2. Excluding 2006-2008 data is not an option either because it will result in loss of part of the stress period performance. This leaves CAMU at no choice but to carve out a middle period as OOT hold out sample to follow MRM's OOT requirement.

3. To ensure the parameter stability and OOT hold out does not cause significant information loss, CAMU further conducted parameter stability analysis and revise/dropped instable parameters. The parameter stability analysis was two-folds including

   a. CAMU estimates the same specification on the 20% OOS sample to ensure consistent parameter sign, significance and reasonable magnitude

   b. CAMU further estimates the same specification on the entire data including the OOT period to ensure consistent parameter sign, significance and reasonable magnitude

It is worth noting though, as CAMU expected, holding out OOT data could potentially result in loss of performance information, especially given that the Jan2012-March2014 period represented the initial phase of economic recovery with high volume of various types of loan modifications and refinancing due to extremely low interest rate and foreclosure settlement(2012) etc. However, given the model's strong performance (backtest) for the OOT period, it can be concluded that the impact from such exclusion is expected to be small as the model development data still comprised of sufficient performance data covering both economic boom and bust. A few transitions that are heavily impacted by portfolio management policy over time, such as IVP transitions, are always statistically challenging to model in the first place. For additional details on the OOT backtest, please refer to Section 6.3 of the MDD. Also please be aware that the OOT period exclusion only pertains to the performances observed during this OOT period. Loans that originated during this period and stayed active behind the OOT period, are included as part of model development sample.

Justification for not including pre-2008 data in LGD model development sample- As per Table 4.1.3.4 pertaining to the LGD Exclusion Waterfall; exclusion # 10 excludes all loans prior to 2008. This was done as the historical loss performance during the earlier development period of the DataMart (pre-2008) exhibited high discrepancy with finance system mainly attributed to lack of a reconciliation process which was establish and enhanced early in 2009 between Risk DataMart and Finance. Based on a deep dive analysis that was conducted in collaboration with the REL DataMart team, it was revealed several discrepancies within the Risk Loss data prior to 2008, attributable to either manual entry errors or top line adjustments to actual loss numbers. Since the reconciliation error between Risk DataMart's Loss file and PEARL (master data file) for the pre-2008 loss data exceeded well over the 1.5% threshold set by REL DataMart team, it was considered prudent to not use the pre-2008 data as part of the Severity model development. For additional details, please look at the attachment – '1.1 Model development data sources and Mortgage Transformation considerations' for pertinent details.

2. MRM Question - In addition to the above points, Sponsors are required to provide rationale/supporting evidence on the below details:

- Rationale for not considering the long sample as the development sample in case recent period data is needed to capture the recent trend

Answer – Please note that the model development data includes the recent most period as part of its development data pool. Hence this question is not valid.

- Analysis and evidence how the development data completes an economic cycle which is a GP1 requirement. Please note that Jan 2012- Mar 2014 period was initial phase of economic recovery with high volume of various types of loan modifications and refinancing due to extremely low interest rate, foreclosure settlement in 2012 etc. Dropping this period may not capture the impact of these important events. Also, there was a big change in portfolio composition between Jan 2012– Mar 2014.

Answer – Bank's Portfolio composition continues to evolve with business strategy leading to changes over time. The current model is built based on loan characteristics thereby allowing it to capture responsiveness to loan characteristics.

In addition, changes in the macroeconomic environment also lead to different borrower behaviors over time. One such instance is observed during the selected OOT period for the model, where higher rates of voluntary prepayment (VP) were observed over many months during a phase of economic recovery. As noted above, the model is developed using loan characteristics as well as observed borrower behavior and macroeconomic drivers over a long time span, and as discussed in section 6.3.2.5 (see also attachment 6.3.2.5 Model OOT Backtest.xlsx), the model is able to predict this rising VP trend within the threshold of error (per MRM Testing Guidance).

- Please provide details on how development data captures increasing interest rate period.

Answer – Please see the swap rate trends over time. Swap rates are the rate attributes used in MOD model to factor in the interest rate movements in the economy which includes the increasing rate environment starting 2015Q4.

- Details on how the panel data was created.

Answer - The panel data creation process (which includes the raw data and the subsequent loan exclusions) have been discussed in Section 4.1.3. Please refer to this section for additional details.

- Provide details of the OOT sample to specify whether only new originations or data from seasoned loans were included.

Answer – As stated many times before, the OOT period exclusion only pertains to the performances observed during this OOT period. Loans that originated during this period and stayed active behind the OOT period, are included as part of model development sample.

3. MRM Question - For the PD non-modeled transitions, the model uses the data from Apr2009-Dec2011 for stress roll rate & Apr2014-Dec2017 for non-stress roll rate. Please provide the rationale for this selection

Answer - The Stress period values are volume weighted mean of transition rates based on data from 200904 (Please note that the MOD Program was initiated Q2-2009 and data on MODs became available starting 2009Q4). US HPI reaches its peak in the second quarter of 2006 and started dropping over time. It hit the bottom in 2009 and remained low until the first quarter of 2012. Instead of using the exact 27month stress period defined by MRM, the stress value selection covered this entire period of HPI decrease to avoid over-fitting the model performance due to data volatility or at odd with MRM's "blind back test" requirement.

4.1.4 Data Sampling Methodology and Results

[If the model is based on a sample, describe the data sampling methodology/scheme (e.g. stratified random sample based on defined strata, etc.), sampling results, sample validation techniques (e.g. out-of-sample/ hold-out sample, out-of-time, and cross-sample validation). Provide evidence that the chosen sample is representative of the total population. In addition to describing the sampling methodology used, justify its appropriateness by comparison of the total population vs sampled population (for e.g. population distribution on key risk drivers, descriptive statistics etc.)].

Modelers commonly use sampling strategy to reduce extremely large sample sizes and improve computational efficiency in the modeling procedures. For the 2019 Method A RM PD model development, modelers used sampling strategy to create a sample for the following reference transition- BUK01->BUK01. On the other hand, observations collected on all other transitions were fully sampled in to the model development process.

All rationale behind the selection of the appropriate transition bucket(s) for the sampling process, the selection of relevant weights and the comparison of key risk drivers before and after sampling have been thoroughly discussed as evidenced below.

The main reasons why BUK01->BUK01 was selected for sampling down because 1) this specific bucket has extremely large sample size (Table 4.1.4.1 shows 65,475,905 loans in BUK01->BUK01); hence sampling down this transition had helped with improving the model efficiency significantly; 2) BUK01->BUK01 is the most common and benign transition used as reference cell in respective logistic regressions. Sampling down this specific transition was expected to have minimal impact to the information enrichment and completeness of potential risk drivers in subsequent modeling.

Based on the rationales stated above, the BUK01-> BUK01 was randomly sampled down by 2%. The sampling rate was so chosen to ensure that the final sample had an adequate number of observations to correctly draw all inferences. Based on existing literature review and research conducted on the prevailing industry trends, the preferred approach suggests that sampling should not adversely affect the power of the statistical test.

As discussed in section 4.1.3, the sampling process was conducted on the data collected on observation month between Jan 2005 and Dec 2017 that has gone through the initial data waterfall process. As shown in table 3,4,5 in section 4.1.3, before the 2% sampling down process, BUK01->BUK01 transition has 65,475,905 records and the total number of observations including all transitions was 77,083, 574. After sampling, the numbers reduced to 1,307,971 in the BUK01->BUK01 and 12,915,640 in the entire datasets.

To justify the appropriateness of the sampling, modelers have examined the population distributions of key risk drivers before vs. after sampling, where risk drivers include marked to market CLTV, Refresh FICO, current principal balance, HPI 12m change, and Unemployment 12m change. This analysis is shown below in Table 4.1.4.1 for BUK01->BUK01 transition, which clearly demonstrates that for the BUK01->BUK01 population, the distributions of all the key risk drivers are very comparable before vs. after sampling. This analysis has provided solid evidence that the sampling have no significant impact to the key risk drivers as model inputs.

Table 4.1.4.1: Comparison of Major Risk Drivers before and after Sampling

Variable	Prior Sample down (N=39,372,452)	Percent	Post Sample down w/ weight (N=39,338,520)	Percent	Absolute Percent Diff
mtm_CLTV (market-to_market CLTV)					
0 - 60	6,441,935	16.36%	6,444,120	16.38%	0.02%
60 - 80	7,429,053	18.87%	7,422,480	18.87%	0.00%
80 - 90	6,602,363	16.77%	6,603,040	16.79%	0.02%
90 - 100	7,781,497	19.76%	7,749,440	19.70%	0.06%
> 100	11,117,604	28.24%	11,119,440	28.27%	0.03%
S_M_FicoRefresh (Refresh Fico)					
<350	0	0%	0	0%	0.00%
350 - 600	2,569,540	6.53%	2,552,200	6.49%	0.04%
600 - 700	11,156,406	28.34%	11,161,720	28.37%	0.03%
700 - 750	11,743,273	29.83%	11,769,680	29.92%	0.09%
750 - 850	13,903,233	35.31%	13,854,920	35.22%	0.09%
D_M_PRIN_BAL (Principle Balance)					
< 80K	33,859,744	86%	33,841,240	86.03%	0.03%
80K - 120K	3,247,687	8.25%	3,239,920	8.24%	0.01%
120K - 200K	1,640,176	4.17%	1,625,240	4.13%	0.04%
200K - 250K	278,402	0.71%	279,880	0.71%	0.00%
250K - 417K	311,583	0.79%	314,760	0.80%	0.01%
> 417K+	34,860	0.09%	37,480	0.10%	0.01%
P_M_Hpi12MChange_LP (HPI 12m Change)					
<-5.19	10,966,392	27.85%	10,967,920	27.88%	0.03%
-5.19- -0.15	10,145,886	25.77%	10,125,320	25.74%	0.03%
-0.15 ~ 4.73	9,587,961	24.35%	9,581,200	24.36%	0.01%
> 4.73	8,672,213	22.03%	8,664,080	22.02%	0.01%
R_M_State_UnempB12M (Unemployment 12m Change)					
< 0.89	9,628,250	24.45%	9,603,240	24.41%	0.04%
0.89 - 0.96	8,488,348	21.56%	8,479,000	21.55%	0.01%
0.96 - 1.14	10,096,527	25.64%	10,098,040	25.67%	0.03%
>1.14	11,159,327	28.34%	11,158,240	28.36%	0.02%

Sampling strategy was not applicable for the 2019 Method A RM Mortgage LGD model.

4.1.5 **Data cleaning and preparation**

What were the data cleaning steps?

[Describe outlier treatment, observation removal, missing data treatment.

Also describe remediation of any data quality issues.]

In the 2019 Method A RM PD and LGD modeling process, prudent and thorough data checks have been conducted to confirm overall sound data quality, as already described in section 4.1.1 to 4.1.3. Based on the observations of scrutinized data reconciliation / quality check reports and descriptive statistics analysis, data cleaning treatments were performed to mitigate the impact of missing values to ensure continuity in the data. Treatments include the following considerations.

1. For the continuous variables except for FICO, missing values were imputed by the median value of the data that has gone through the initial waterfall process and a corresponding missing control variable was created. In particular, considering the significant differences between the CMI and CPB portfolio in terms of property value, loan size and borrower credit profile, missing value on CMI and CPB loans was imputed by their respective portfolio median.

   A relatively high missing rate was observed on Refresh FICO (S_M_FICORefresh), particularly among the loan just booked into the portfolio. So for the newly originated loans (month on book<=3), if the refresh FICO was missing, CAMU used decision FICO to feed in the value. Otherwise the portfolio median was used for the missing value imputation and in this case a missing control variable would be created.

   The use of a missing control variable means the choice of imputation default value had no impact on the model. Any combination of the missing control variable and default value had the same result. Missing values are always imputed instead of deleting the records with missing values from the development data, with the benefits that it can retain the maximum available data coverage, and more importantly, it keeps the potential unique prediction power of the missing values in the modeling. Second, the missing values are imputed by the median, because the median is deemed as the "neutral" value of the valid range, which is not affected by large outliers. Third, using missing indicator as control variable allows modelers to accurately model the marginal effect of the missing values for the variable independently from the effect of its valid values, and differentiate the true mean values records from the imputed records, which also serve as reassurance to eliminate potential estimation bias due to missing imputation by the mean.
2. Limited capping methods were used to treat outliers. Cap was chosen as the 99% percentile + 2*STD and was calculated for CMI and CPB respectively.

3. No observation was blindly removed due to missing value alone. Instead, the population exclusion waterfall was fully analyzed and any exclusions were supported with sound rationales (please refer to next section of data exclusion).
4. Within the Severity model, for loans (primarily first lien) with Private Mortgage Insurance (PMI) coverage, if the PMI coverage ratio is missing, it is substituted with a 25% imputed rate. On the entire LGD development data of 304,926 accounts, 1970 carry PMI coverage. PMI coverage percentage ranges from 6% to 50%, Table 4.1.5.1. below illustrates the PMI coverage percentage (PMI_pct) univariate distribution on LGD development data, which shows that 25% is the median value and most common of PMI coverage percentage.

Table 4.1.5.1: PMI Coverage Percentage Univariate Distribution

Portfolio	N Obs	Variable	N	N Miss	Mean	Minimum	1st Pctl	5th Pctl	10th Pctl	25th Pctl	50th Pctl	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Maximum	Std Dev
1st	1970	pmi_pct	1970	0	0.257437	0.06	0.12	0.12	0.12	0.2	0.25	0.3	0.35	0.35	0.35	0.5	0.071738

5. Within the Severity model, for all first mortgage loans with missing balances, these are also adjusted as UPB/0.25. For the first mortgage loans with missing balances, it is imputed with a 25% rate for the LGD development sample. This is based on the empirical analysis of the LGD development data. As shown, the 25% rate is the median value.

Table 4.1.5.2: First Mortgage Balance Univariate Distribution

Variable	Label	N	N Miss	Mean	Minimum	1st Pctl	5th Pctl	10th Pctl	25th Pctl	50th Pctl	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Maximum	Std Dev
FST_MTG_BAL	FST_MTG_BAL	223498	0	267,556	1000	48753.65	82400	104500	151998.5	228000	336000	467998	573744	915284.3	12237482	180317.48
max_brd_prin		223498	0	74,112	2850	12000	19800	25000	36000	57550	90000	143000	185000	317241	9000000	65879.17
max_brd_prin/FST_MTG_BAL =				0.28						0.25						

6. For all junior liens within the LGD development pool, the junior ratio is imputed at 20%,presented below is the empirical evidence .

Table 4.1.5.3: Junior Ratio Univariate Distribution

portfolio='2nd' and excl_all=0 and CPB_ind=0

Analysis Variable : junior_ratio														
N	N Miss	Mean	Minimum	1st Pctl	5th Pctl	10th Pctl	25th Pctl	50th Pctl	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Maximum	Std Dev
236831	0	0.2155767	0.001503	0.06506	0.108047	0.111111	0.165523	0.2	0.231194	0.326148	0.403651	0.570601	0.989289	0.094675

7. For the RM PD model – second lien loans, CAMU assumed a constant junior ratio to derive the balance of the first lien to determine the most updated mark-to-market CLTV of the loan. For most of the second liens, a special challenge to determine the CLTV is the dynamic value of the first lien balance associated with the property is not available in the source data. The only available information is the first lien balance when this second lien was originated. To address the issue, CAMU assumes that the ratio between the first lien and second lien stays constant throughout the life of the second lien.

For example, suppose a second lien was originated with 20K UPB, and at its origination, there was also an 80K UPB first lien associated with the property, then the junior ratio at origination is 20K/80K=25%. Then at any observation month, if the UPB of this second lien drops to 10K, then CAMU assume its first lien balance also proportionally reduces to 10/25%=40K. Then the total of 50K will be used to calculate the combined CLTV on this second lien.

Refer to section 4.1.4, Data Water Fall for RM never-modified Loans for the data cleaning or exclusion. Each model driver that required imputation or treatment has been listed in Table 4.1.5.4. To keep the variables in a reasonable range, CAMU imputed variables on a case-by-case basis. There was no sign of bias by examining the deviation of mean values after imputation from the values before imputation.

Table 4.1.5.4. Method A RM PD Variable Treatment

	Variable	Treatment	% missing before treatment	Median	Maximum before treatment	Maximum after treatment (99th pct+2 StdDev)	Mean before treatment	Mean after treatment
CMI Only	D_M_PRIN_BAL	cap	0%	111,436	13,025,000	1,445,666	181,673	179,915
	P_M_PresIntSpread	cap	0%	3.13	16.34	14.88	3.45	3.45
	P_S_OrigIntSpread	cap	0%	2.39	71.81	12.09	2.67	2.67
	S_M_FicoRefresh	missing value	12%	660	850	850	669	671
	P_M_CURR_NOTE_RATE	missing value, cap	0%	6.63	21.00	18.75	6.89	6.89
	P_S_ORIG_NOTE_RATE	missing value, cap	1%	6.84	76.88	19.16	7.22	7.22
	P_S_BE_DEBT_RATIO	missing value, cap	9%	33.65	995.00	100.00	30.85	30.95
	D_S_ORIG_PROP_AMT	missing value, cap	1%	175,000	32,599,119	2,861,921	308,379	301,717
CPB Only	D_M_PRIN_BAL	cap	0%	693,020	53,400,000	8,136,479	998,808	977,318
	P_M_CURR_NOTE_RATE	cap	0%	4.50	10.88	9.55	4.49	4.49
	P_S_ORIG_NOTE_RATE	cap	0%	4.88	17.00	10.50	4.81	4.81
	P_M_PresIntSpread*	cap	0%	1.94	8.31	8.31	2.03	2.03
	P_S_OrigIntSpread*	cap	0%	1.83	6.97	6.97	1.86	1.86
	S_M_FicoRefresh	missing value	14%	747	850	850	744	775
	P_S_BE_DEBT_RATIO	missing value, cap	2%	29.58	100.00	100.00	29.64	29.64
	D_S_ORIG_PROP_AMT	missing value, cap	0%	1,320,000	153,840,000	18,974,945	2,053,902	2,011,603
All	N_M_age_oldest_trd	missing value, cap	54%	215	983	752	237	225
	N_M_n_open_mort	missing value	54%	1	79	79	1.5	1.24

	Variable	Treatment	% missing before treatment	Median	Maximum before treatment	Maximum after treatment (99th pct+2 StdDev)	Mean before treatment	Mean after treatment
	N_M_Age_Oldest_Mtg	missing value, cap	36%	124	613	482	127	126

Refer to Table 4.1.3.7.: Data Water Fall for Severity Model for the data exclusions pertaining to the Severity model. For all other data cleaning and preparation, any model driver that required imputation or treatment has been listed in Table 4.1.5.5. To keep the variables in a reasonable range, CAMU imputed variables case by case. For instance, principal balance (prin_bal) was capped at 99% percentile + 2*STD to remove the outlier values. There was no sign of bias as can be seen by examining the deviation of mean values after imputation from the values before imputation.

Table 4.1.5.5. LGD Model Variable Treatment

Variable	Reason for Cleaning	Data Cleaning Logic	%Missing	Mean before Imputation	Mean after Imputation
prin_bal	out-of-range	cap at 99%+2STD	0%	$143,744	$142,287
fts_mtg_bal	out-of-range, missing	cap at 99%+2STD, impute with medium value	26%	$201,446	$160,051

Were any data proxies used?

[A data proxy is an alternate variable used to calculate missing values. Also include macroeconomic data proxies.]

Example: *90 day delinquent loans may be used as a proxy for default.*

No data proxies were used.

Was any data excluded for data quality or business reasons? What was the rationale for the exclusion?

[Describe data exclusions due to quality concerns or due to business reasons (e.g., loan modifications, discontinued products). Provide the rationale and explain why the exclusion is reasonable. Provide a waterfall chart based on exclusions and filtering (e.g., with number of observations.)

Also describe if any macroeconomic data was excluded for data quality reasons.]

There was no macroeconomic data exclusions made for quality reasons. For both PD and LGD model exclusions, please refer to Section 4.1.3. The data exclusions stated above in table 4.1.3.1 and table 4.1.3.4 for data waterfall exclusion details for PD and LGD model respectively, were mainly due to out of scope population, not part of the official loan list, different accounting treatment for Government loans and known data quality of the loss table in the early period of its development prior to 2008. The overall exclusion is relatively small and does not represent the current state of the data in the portfolio. In particular, for 180+DPD to VP model, government loans prior to Jan 2008 were excluded from model development because abnormally high 180+DPD to VP rate, which is 4.5%, were observed in the data during that period.

What data discrepancies were observed and what actions were taken to mitigate them? What was the impact of data discrepancies, cleaning, data exclusions and remediation?

[Describe each significant discrepancy identified during data quality checks and how it was treated. Include missing data (e.g., missing FICO scores).

Describe how the treatment affected the portfolio. Comment on changes in GCL/ENR, number of accounts, etc. and potential impact on the model. Quantify the impact.

It must be clear whether the data can still be used for model development, and that the remediation was conservative and prudent. Also include the impact of data exclusions.]

PD Development data

During the course of 2019 Method A RM model development, as part of the rigorous data quality checks and descriptive statistical analyses, few spikes were observed from the one-month-ahead roll rate time curve. Such spikes indicate the natural intervention of business actions taken at that time. Therefore the spikes were documented and actively communicated with business users to seek policy explanations.

There was no action taken in modeling practice for these spikes for two reasons. First, these spikes occurred infrequently in the performance history data. So they are expected to have minimal impact on model forecasting. Second, these spikes may represent valid data points at various periods. Therefore, it is preferable to keep them intact in the modeling data.

1) Spike in roll rate from BUK7 to BUK 2,3,4,5,6 in June'08– During the RM PD model development process, ,two spikes were noted in the partial cure roll rates from BUK7 to BUKs2-6. These spikes could be attributable to the debt repayment plans that were put in place for the bankruptcies claimed under

Chapters 13/11, loan modifications and subsequent payment adjustments made. These spikes affected the Method A -Residential Mortgage first lien delinquencies.

2) Spike in roll rate from BUK6, 7 & BUKs 2,3,4,5,6 in Feb'10 –Spikes were noted in the partial cure roll rates from Buk7,6 to Buks2-6. These spikes, as before were attributable to the payment adjustments made on these loans

Table 4.1.5.1: BUK7 -> Partial Cure BUKs 2-6: Two spikes in 200806 and 201002

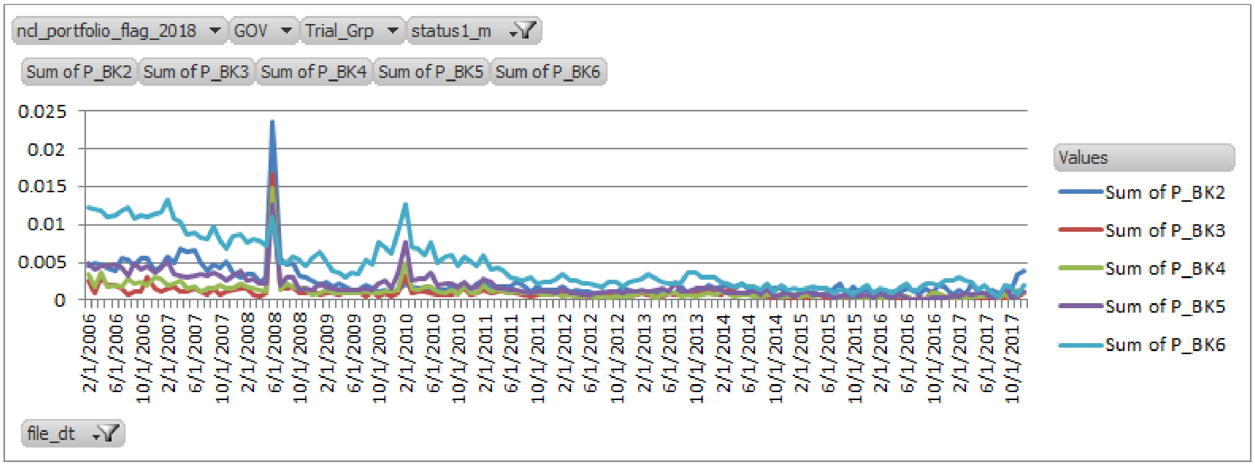
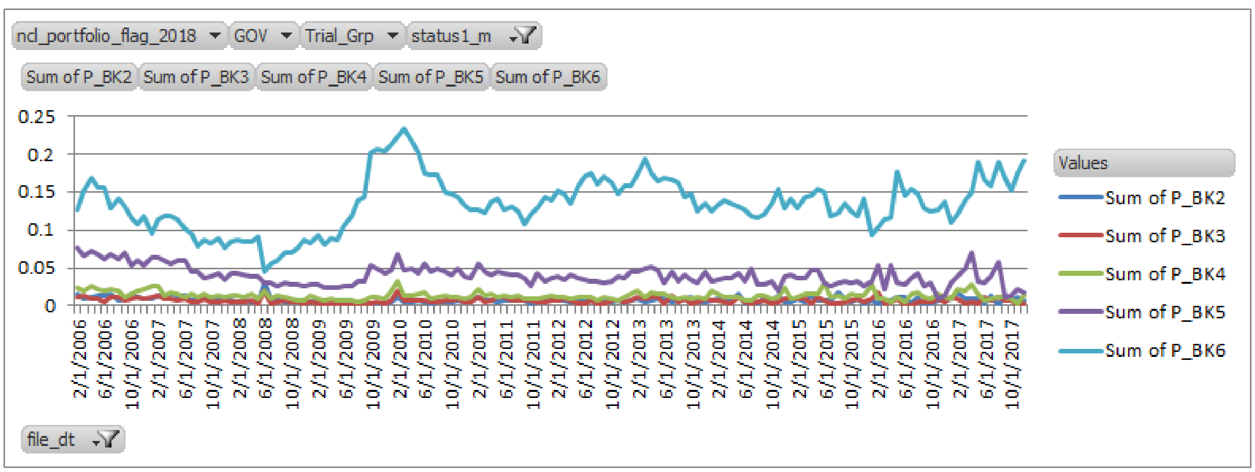


Table 4.1.5.2: BUK6 -> Partial Cure BkT2-6: one spike in 201002



3) Increase in GOV IVP Rate since 2016 - A Government loans is typically characterized having 97 LTV at origination with corresponding high DTI (Debt-to-income) ratio. Operationally the business has always faced challenges in government loan liquidation as the process to liquidate government loans is not the same as conventional loans. Government loans are insured, and in order to file the insurance claim and obtain the insurance proceeds the property must be conveyed to HUD. As part of this process, Bank must meet HUD's guidelines for conveyance. These guidelines require eviction to be complete, as well as for the property to be in "conveyance condition." This is a higher standard than what is required for conventional loans, where there are no similar requirements around eviction and property condition.

For conventional loans elements such as occupancy and property condition impact REO price but are not preconditions for a loan to be sold through REO. These problems in conveying government loans are caused by two factors: 1) changing and inconsistent interpretations of HUD requirements, impacting all servicers, and 2) idiosyncratic challenges Bank has faced in the management of HUD conveyance where the operational group has been transitioned between sites and has undergone multiple leadership changes. Policy wise, when a government loan becomes IVP and goes into subsequent foreclosure sale; and if BankMortgage wins the bid during the foreclosure induced sale; this loan would tend to remain active on the Bank's system (data warehouse) until conveyed to HUD (Dept. of Housing & Urban Development). But the loan even though showing 'active' status is not a real loan as the property was already foreclosed and the collateral is owned by BankMortgage. This observable data trend with the uptick in the IVP rates for Government loans, contradicts with the prevailing macro-economic trends in recent time periods which are marked with significant and consistent home price appreciation and unprecedentedly low unemployment rates. On a strategic note, as mentioned earlier, there were a transfer of middle office functionalities, who managed the foreclosure of the government loans, which lead to instability in the management process and the subsequent issuance of MRA's from OCC around the management of debts previously contracted (DPC). Although CMI had increased performance since then, especially in regards to the increased management and regulatory scrutiny in this area, particularly related to debts previously contracted (DPC) exposure, this also has played a role in the uptick of the IVP rates.

Table 4.1.5.3: Trend in GOV IVP Rate

LGD Development data

For LGD Model development data, the relevant population exclusions are discussed above in the waterfall logic. In particular, the business maintains a one-timer loan level list that lists all loans with abnormal losses. For additional information on one-timer events and treatments, please refer to attachment "3.2_Memorandum_One-Timer Exclusions - 111417.docx". Since these losses do not conform to business rationale or economic intuition, all such loans are excluded from the development data. The one timer list also includes loss spike(s)/dip(s) that occur due to strategic business intervention decision.

4.1.6 Data transformation

[Describe any statistical systematic transformations applied to the performance, segmentation, account or macroeconomic data, e.g., expansion/mapping of national (or international regional) macro drivers to state-level (or country level) macro drivers, lags, differencing related to stationarity or seasonality.]

What is the rationale for performance, segmentation and account data transformation?

The 2019 Method A RM PD and LGD variable transformations include all the new variable transformations that were conducted this modeling cycle based on economic expectations and comprehensive statistical analyses during the variable selection process. These variable transformations were performed to improve model performance particularly on granular segments, improve the model's sensitivity to certain key risk drivers, capture the non-linearity of some specific trends, address previous feedbacks from model reviewers / regulators and business users and lastly improve the model's overall performance. The detailed economic rationale and related analyses for these variable transformations are further discussed in section 5.1.4 of variable selection.

The variable transformations that were attempted this modeling cycle on this specific model suite included the following –

- Adding Caps and/or Floors – Caps and Floors sets upper and lower bounds to the values that the numerical attributes can take. Caps and Floors are primarily done to align the model's performance with business intuition. A good example to illustrate here would be usage of the floor for the marked-to-market LTV in the Severity model. For added rationale on the use of the floor at MTM LTV = 100, please refer to Section 3.1. of the MDD.

- Introducing Interaction Effects- - Interaction variables were created in the model suite to account for the differential credit risk effects between the stress / non-stress cycles. Another reason of interaction variables is to model the differentiated sensitivities to risk drivers by key sub-groups without separate segmentation. For example, within the RM PD model BUK01->BUK2 transition, the state level unemployment rate was interacted with the CMI portfolio, whereas the unemployment 12 month ago was interacted with the CPB portfolio indicator, as it's observed that historically, the CPB portfolio demonstrated a delayed response to the change in macro-economic environment along with a different magnitude of sensitivity.

Spline Transformations- Spline transformed variables are transformations that are used to capture non-linear relationships of the root variables, the choices of spline knots are based on business / economic intuitions and validated by bivariate analysis as visualized in bubble plots. For example within the Severity stage two model a spline is created for the marked-to-market CLTV at 90 which aligns with the business policy requirements that changes for CLTV 90 and over(Figure 4.1.6.2). For the RM PD model, as shown in the bubble plot below (Figure 4.1.6.1), BUK01->BUK2 rate increases with CLTV when CLTV is greater than 40, but the slope is less steep when CLTV is above 80. In this case, mark-to-market CLTV splines were created at 40 and 80 to capture the differentiated sensitivities of the delinquency rate changes to a unit change of CLTV within different ranges of CLTV.

Figure 4.1.6.1 – Spline Transformation Justification for the RM PD Model – Mark-to-Market CLTV

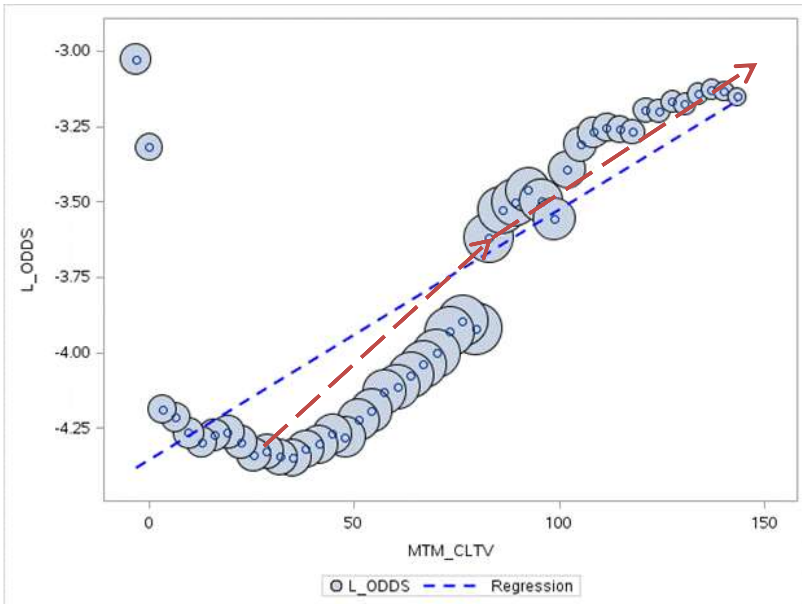
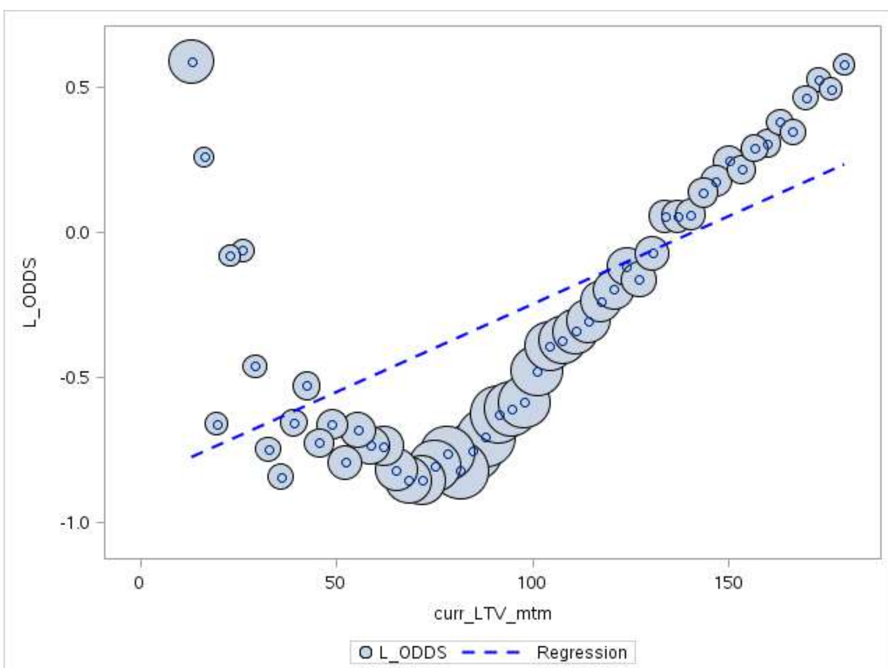
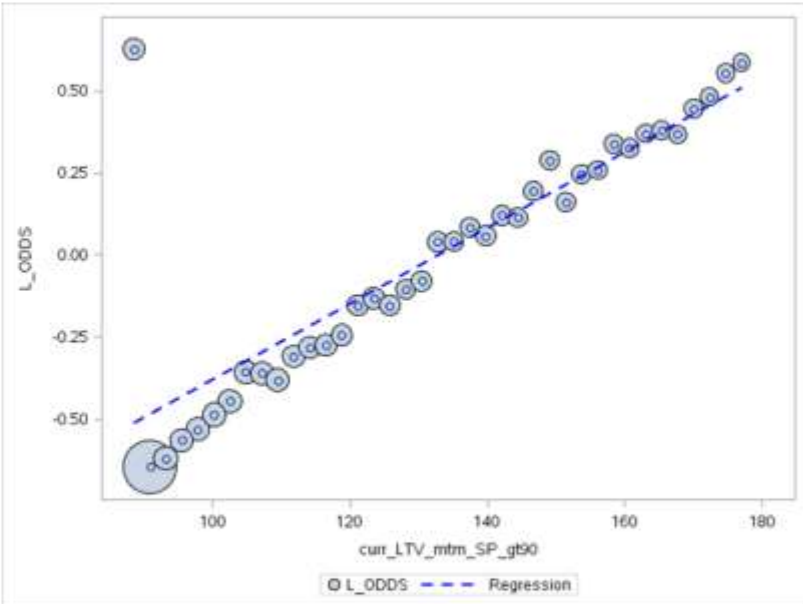


Figure 4.1.6.2 – Spline Transformation Justification for the LGD Model – Mark-to-Market CLTV

Before Introduction of Splines



After Introduction of Splines (Monotonicity is maintained)



Presented below are the variable transformations attempted on the 2019 Method A RM PD and Severity models respectively.

Table 4.1.6.3: 2019 Residential Mortgage PD Variable Transformation

Transform Varaiable (Spline)	Raw Variable	Meaning	Source field (or formula if calculated field)
D_M_PRIN_BAL_10k_SP5LSP15_CMI	PRIN_BAL	spline of principle balance interacts w ith cmi indicator. Spline sets as betw een 50K and 150K.	D_M_PRIN_BAL_in10k_SP5LSP15 = min(max(D_M_PRIN_BAL_in10k,5),15); D_M_PRIN_BAL_10k_SP5LSP15_CMI = D_M_PRIN_BAL_in10k_SP5LSP15*(cpb_ind=0);
D_M_PRIN_BAL_in10k_LSP10	PRIN_BAL	spine of principle balance <= 100K	D_M_PRIN_BAL_in10K_LSP10=min(D_M_PRIN_BAL_in10K,10)
D_M_PRIN_BAL_in10k_LSP20_CMICONV	PRIN_BAL	spine of principle balance interacts w ith CMI and non-GOV indicators. Spline sets as <= 200K.	D_M_PRIN_BAL_in10k_LSP20 = min(D_M_PRIN_BAL_in10k,20); D_M_PRIN_BAL_in10k_LSP20_CMI = D_M_PRIN_BAL_in10k_LSP20*(cpb_ind=0); D_M_PRIN_BAL_in10k_LSP20_CMICONV = D_M_PRIN_BAL_in10k_LSP20_CMI*(GOV=0);
D_M_PRIN_BAL_in10k_SP10_GOV	PRIN_BAL	spine of principle balance interacts w ith GOV indicators. Spline sets as >=100K.	D_M_PRIN_BAL_in10K_SP10=max(D_M_PRIN_BAL_in10K,10); D_M_PRIN_BAL_in10K_SP10_GOV = D_M_PRIN_BAL_in10K_SP10*(GOV=0);
D_M_PRIN_BAL_in10k_SP20_CMICONV	PRIN_BAL	spine of principle balance interacts w ith CMI and non-GOV indicators. Spline sets as >= 200K.	D_M_PRIN_BAL_in10k_SP20_CMICONV = D_M_PRIN_BAL_in10k_SP20_CMI**(GOV=0);
D_M_PRIN_BAL_SP100K_CMI_1st	PRIN_BAL	spline of principle balance interacts w ith CMI and 1st lien indicators. Spline sets as >=100K.	D_M_PRIN_BAL_SP100K_CMI = D_M_PRIN_BAL_SP100K*(cpb_ind=0); D_M_PRIN_BAL_SP100K_CMI_1st =D_M_PRIN_BAL_SP100K_CMI*(B_S_Lien2nd_IND=0);
D_M_PRIN_BAL_SP100K_CMI_HPIDecr	PRIN_BAL	spline of principle balance interacts w ith CMI indicator and HPI indicators. Spline sets to >=100K.	D_M_PRIN_BAL_SP100K_CMI_HPIDecr =D_M_PRIN_BAL_SP100K_CMI*(HPI_tm12_ratio<=1);
HPI_APP_LSP1_CMI			
HPI_tm12_ratio_LSP1	HPI	HPI index in past 12 mos. Spline set to <= 1.	HPI_tm12_ratio_LSP1=min(HPI_tm12_ratio,1);
HPI_tm12_ratio_LSP1_CMI_CONV	HPI	HPI index spline interacts w ith CMI and non-GOV indicators. Spline sets to <= 1.	HPI_tm12_ratio_LSP1_CMI= HPI_tm12_ratio_LSP1*(CPB_IND=0); HPI_tm12_ratio_LSP1_CMI_CONV = HPI_tm12_ratio_LSP1_CMI*(GOV=0);
HPI_tm12_ratio_SP1	HPI	HPI index in past 12 mos. Spline set to >=1.	HPI_tm12_ratio_SP1=max(HPI_tm12_ratio,1);
HPI_tm12_ratio_SP1_2nd	HPI	HPI index in past 12 mos interacts w ith 2nd lien indicator. Spline set to >=1.	HPI_tm12_ratio_SP1_2nd = HPI_tm12_ratio_SP1*(B_S_Lien2nd_IND=1);
HPI_tm12_ratio_SP1_MOBGT120	HPI	HPI spline interacts w ith MOB spline.	HPI_tm12_ratio_SP1_MOBGT120=HPI_tm12_ratio_SP1*N_M_MOB_GT120;
income_12m_ratio_LSP1	INCOME	income ratio in past 12 mos. Spline sets to <= 1.	income_12m_ratio_LSP1=min(income_12m_ratio,1);
INX_SP500_12m_ratio_LSP1	INX_SP500	SP500 ratio in past 12 mos. Spline sets to <=1.	INX_SP500_12m_ratio_LSP1 = min(INX_SP500_12m_ratio,1);
INX_SP500_12m_ratio_SP1	INX_SP500	SP500 ratio in past 12 mos. Spline sets to >=1.	INX_SP500_12m_ratio_SP1=max(INX_SP500_12m_ratio,1);
N_M_Curr_lt_Orig_Rate_LSP48	CURR_NOTE_RATE, sw ap	Current note rate < rate at origination. Spline set to <=48.	N_M_Curr_lt_Orig_Rate_LSP48 = min(N_M_Curr_lt_Orig_Rate,48);
P_M_PresIntSpread_LSP3pt5_CMI	curr_note_rate, sw ap	interest spread interacts w ith CMI indicator. Spline set to <= 3.5	P_M_PresIntSpread_LSP3pt5 = min(P_M_PresIntSpread,3.5); P_M_PresIntSpread_LSP3pt5_CMI = P_M_PresIntSpread_LSP3pt5*(cpb_ind =0);
P_M_PresIntSpread_LSP3pt5_FRM	curr_note_rate, sw ap	interest spread interacts w ith fixed rate indicator. Spline sets to <= 3.5	P_M_PresIntSpread_LSP3pt5_FRM = P_M_PresIntSpread_LSP3pt5*B_S_FIXED_RATE_IND;
P_M_PresIntSpread_LSP5_CPB	curr_note_rate, sw ap	interest spread interacts w ith CPB indicator. Spline sets to <= 5	P_M_PresIntSpread_LSP5 = min(P_M_PresIntSpread,5); P_M_PresIntSpread_LSP5_CPB = P_M_PresIntSpread_LSP5*cpb_ind;
P_M_PresIntSpread_LSP5_PRA	curr_note_rate, sw ap	interest spread interacts w ith reset indicator. Spline sets to <= 5	P_M_PresIntSpread_LSP5_PRA = P_M_PresIntSpread_LSP5*(B_S_PreResetArm_IND);
P_M_PresIntSpread_LSP5_Pre2010	curr_note_rate, sw ap	interest spread interacts w ith pre2010 indicator. Spline sets to <= 5	P_M_PresIntSpread_LSP5_Pre2010 = P_M_PresIntSpread_LSP5*(Post2010_Orig=0);

N_M_logHpiChange_Orig_LSP0	HPI, LOAN_ORIG_DT	Log transformation HPI change associated with the origination. Spline sets to <= 0.	N_M_logHpiChange_Orig_LSP0=min(N_M_logHpiChange_Orig,0);
N_M_logHpiChange_Orig_SP0		Log transformation HPI change associated with the origination. Spline sets to >= 0.	N_M_logHpiChange_Orig_SP0=max(N_M_logHpiChange_Orig,0);
MTM_CLTV_LSP100	prin_bal, orig_prop_amt, HPI,fst_mtg_bal, origBalance, origCLTV, origLTV	Market-to-market CLTV spline <= 100	MTM_CLTV_LSP100=min(MTM_CLTV,100);
MTM_CLTV_LSP60		Market-to-market CLTV spline <= 60	MTM_CLTV_LSP60=min(MTM_CLTV,60);
MTM_CLTV_SP40_1st_CONV		Market-to-market CLTV interacts with non-GOV indicator. Spline sets to >=40	MTM_CLTV_SP40=max(MTM_CLTV,40); MTM_CLTV_SP40_1st = MTM_CLTV_SP40 * B_S_lien1st_IND; MTM_CLTV_SP40_1st_CONV= MTM_CLTV_SP40_1st*(GOV=0);
MTM_CLTV_SP40LSP120		Spline of MTM CLTV with values set to between 40 and 120.	MTM_CLTV_SP40LSP120=min(max(MTM_CLTV,40),120);
MTM_CLTV_SP40LSP120_GOV		MTM CLTV spline interacts with GOV indicator. Spline set between 40 and 120.	MTM_CLTV_SP40LSP120_GOV=MTM_CLTV_SP40LSP120*GOV
MTM_CLTV_SP60		MTM CLTV spline >= 60	MTM_CLTV_SP60=max(MTM_CLTV,60);
MTM_CLTV_SP80		MTM CLTV spline >= 80	MTM_CLTV_SP80=max(MTM_CLTV,80);
MTM_CLTV_SP80_2nd_CONV		MTM CLTV spline interacts with 2nd lien and non-GOV indicators. Spline sets to >=80	MTM_CLTV_SP80_2nd_CONV= MTM_CLTV_SP80_2nd*(GOV=0);
MTM_CLTV_SP80_1st_CONV		MTM CLTV spline interacts with 1st lien and non-GOV indicators. Spline sets to >=80	MTM_CLTV_SP80_1st = MTM_CLTV_SP80 * B_S_lien1st_IND MTM_CLTV_SP80_1st_CONV= MTM_CLTV_SP80_1st*(GOV=0);
MTM_CLTV_SP40_1st_CONV		MTM CLTV spline interacts with 1st lien and non-GOV indicators. Spline sets to >=40	MTM_CLTV_SP40=max(MTM_CLTV,40); MTM_CLTV_SP40_1st = MTM_CLTV_SP40 * B_S_lien1st_IND; MTM_CLTV_SP40_1st_CONV= MTM_CLTV_SP40_1st*(GOV=0);
MTM_CLTV_GT100_GOV		MTM CLTV spline interacts with GOV indicator. Spline set >= 100	MTM_CLTV_GT100_GOV=(MTM_CLTV>100)*(GOV=1);
MTM_CLTV_SP60LSP80_CPB		MTM CLTV spline interacts with CPB indicator. Spline set between 60 and 80.	MTM_CLTV_SP60LSP80_CPB = MTM_CLTV_SP60LSP80*cpb_ind;
MTM_CLTV_SP60_HPIDcr		MTM CLTV spline interacts with HPI ratio. Spline value set >=60	MTM_CLTV_SP60_HPIDcr = MTM_CLTV_SP60*(HPI_tm12_ratio<=1);
MTM_CLTV_MOBGT120		MTM CLTV interacts with MOB >120 spline	MTM_CLTV_MOBGT120=MTM_CLTV*N_M_MOB_GT120;
N_M_MOB_LSP36	FILE_DT, LOAN_ORIG_DT	MOB spline <= 36	N_M_MOB_LSP36=min(N_M_MOB_cap,36);
N_M_MOB_LSP36_GOV			N_M_MOB_LSP36_GOV=N_M_MOB_LSP36*GOV;
N_M_MOB_LSP72		MOB spline <= 72	N_M_MOB_LSP72=min(N_M_MOB_cap,72);
N_M_MOB_SP36		MOB spline >= 36	N_M_MOB_SP36=max(N_M_MOB_cap,36);
N_M_MOB_SP36LSP180		MOB spline between 36 and 180	N_M_MOB_SP36LSP180=min(max(N_M_MOB,36),180);
R_M_State_UnempB12M_SP1	unemployment	unemployeemnt in past 12 mos. Spline sets to >=1	R_M_State_UnempB12M_SP1 = max(R_M_State_UnempB12M,1);
R_M_State_UnempB12M_SP1_MOBGT120	unemployment, loan_orig_dt, file_dt	spline of unemployment >=1 interacts with MOB spline >120	R_M_State_UnempB12M_SP1_MOBGT120=R_M_State_UnempB12M_SP1*N_M_MOB_GT120;
S_M_FicoRefresh_LSP640	bacon50_score	refresh fico spline <= 640	S_M_FicoRefresh_LSP640=min(S_M_FicoRefresh,640);
S_M_FicoRefresh_SP640	bacon50_score	refresh fico spline >= 640	S_M_FicoRefresh_SP640=max(S_M_FicoRefresh,640);
S_M_FicoRefresh_SP720	bacon50_score	refresh fico spline >= 720	S_M_FicoRefresh_SP720=max(S_M_FicoRefresh,720);

Table 4.1.6.4: 2019 LGD Variable Transformation

Model	Transform Varaiable	Raw Variable	Meaning	Source field (or formula if calculated field)

2nd lien - state 2: partial loss rate	cLTV_MTM_SP_gt80_le120	PRIN_BAL, FST_MTG_BAL, ORIG_PROP_AMT, HPIV4	Market-to-market CLTV capped at 80	if orig_prop_amt>0 then ORIG_PROP_AMT2=min(ORIG_PROP_AMT_cap,max(1,ORIG_PROP_AMT));*floor at 1,cap at 99%+2STD;
				else orig_prop_amt2=brd_amt2;
				PRIN_BAL2=min(PRIN_BAL_cap,PRIN_BAL);*cap at 99%+2STD;
				if HPI_orig>0 then PROP_AMT_MTM=ORIG_PROP_AMT2*(HPI_inact/HPI_orig);
				curr_LTV_mtm = 100*PRIN_BAL2/PROP_AMT_MTM;
				curr_LTV_mtm=max(min(curr_LTV_mtm,300),0);
				curr_cLTV_mtm=curr_LTV_mtm;
				if portfolio in ('2nd','VA') then do;
				
				curr_cLTV_distress=100*(PRIN_BAL2+FST_MTG_BAL2)/asis3;
				
				curr_cLTV_distress2=100*(PRIN_BAL2+FST_MTG_BAL2)/asis_MTM;
				if HPI_orig>0 and ORIG_PROP_AMT2>0 then curr_cLTV_mtm = 100*(PRIN_BAL2+FST_MTG_BAL2)/(ORIG_PROP_AMT2*(HPI_inact/HPI_orig));
				else curr_cLTV_mtm=curr_cLTV_distress;
				end;
				curr_cLTV_mtm=max(min(curr_cLTV_mtm,300),0);
				cLTV_MTM_SP_gt80_le120=max(min(curr_cLTV_MTM,120),80);

2nd lien - stage 1: zero vs partial	cLTV_MTM_unemp_down	Unemployment, PRIN_BAL, FST_MTG_BAL, ORIG_PROP_AMT, HPIV4	Interaction of market-to-market CLTV with unemployment indicator	Unemp_down_ind=ifn(R_M_UnempB12M<=1,1,0); cLTV_MTM_unemp_down=cLTV_MTM*unemp_down_ind;

2nd lien - stage 1: full vs partial	junior_ratio_SP_lep20	brd_amt, fst_mtg_bal	Spline junior ratio floored at 0.2	junior_ratio=brd_amt2/(fst_mtg_bal2+brd_amt2);
				if junior_ratio<=0 then junior_ratio=0.2;*medium value;
				junior_ratio_SP_lep20=min(junior_ratio,0.2);

1st lien - stage 2: partial loss rate	log_curr_bal	prin_bal	log transformation	curr_bal=PRIN_BAL2; log_curr_bal=log(curr_bal);"

2nd lien - stage 1: full vs partial	log_curr_bal_SP_gt10k_le150k	prin_bal	log principal balance floored at 10K and capped at 150K	PRIN_BAL2=min(PRIN_BAL_cap,PRIN_BAL);*cap at 99%+2STD;
				curr_bal=PRIN_BAL2;
				log_curr_bal_SP_gt10k_le150k=log(min(max(curr_bal,10000),150000));

Model	Transform Varaiable	Raw Variable	Meaning	Source field (or formula if calculated field)
2nd lien - stage 1: full vs partial	log_curr_bal_SP_gt150k	prin_bal	log principal balance cappted at 150K	log_curr_bal_SP_gt150k=log(max(curr_bal,150000));
1st lien - stage 1: Full vs partial	log_curr_bal_SP_le300k	prin_bal	log principal balance floored at 300K	log_curr_bal_SP_le300k=log(min(curr_bal,300000));
1st lien - stage 1: Full vs partial	log_curr_bal_SP_gt300k	prin_bal	log principal balance capped at 300K	log_curr_bal_SP_gt300k=log(max(curr_bal,300000));
2nd lien - state 2: partial loss rate	log_fst_mtg_bal	fst_mtg_bal	log transformation	log_fst_mtg_bal = log(fst_mtg_bal);
1st lien - stage 2: partial loss rate	log_HPI_t_orig_rati	HPI	log transformation	if HPI_orig>0 then HPI_t_orig_ratio=HPI_inact/HPI_orig; log_HPI_t_orig_ratio=log(HPI_t_orig_ratio);
1st lien - stage 1: zero vs partial, 2nd lien - stage 1: full vs partial	log_HPI_tm12_ratio	HPI	log transformation	HPI_tm12_ratio=HPI_inact/HPI_tm12; log_HPI_tm12_ratio=log(HPI_tm12_ratio);
1st lien - stage 1: Full vs partial	log_HPI_tm12_ratio_SP_le0	HPI	log of HPI ratio is past 12 mos floored at 0	log_HPI_tm12_ratio_SP_le0=min(log_HPI_tm12_ratio,0);
1st lien - stage 2: partial loss rate	LTV_mtm_SP_gt90	PRIN_BAL, FST_MTG_BAL, ORIG_PROP_AMT, HPIV4	Market-to-market CLTV capped at 90	LTV_mtm_SP_gt90=max(curr_LTV_mtm,90);
2nd lien - stage 1: full vs partial	R_M_UnempB12M_SP_gt1_le1p2	unemployment	Unemployment of past 12 mos floored at 1 and capped at 1.2	R_M_UnempB12M_SP_gt1_le1p2=min(max(R_M_UnempB12M,1),1.2);
	UnempRate_SP_gt	Unemployment	Unemployment capped at 9	UnempRate_SP_gt9=max(UnempRate,9);
1st lien - stage 1: zero vs partial	LTV_mtm_SP_gt20_le100	PRIN_BAL, FST_MTG_BAL, ORIG_PROP_AMT, HPIV4	Market-to-market CLTV capped at 100 and floored at 20	LTV_mtm_SP_gt20_le100=max(min(curr_LTV_mtm,100),20);

Does the approach to macroeconomic variable selection or transformation significantly differ from approaches in similar countries in the region?

The U.S. real estate market is both mature and idiosyncratic. The approaches taken in nearby countries were not considered.

Macroeconomic variables such as HPI, unemployment rate, interest rate, income, GDP are typical inputs to loss forecasting models. Usage of such variables are often recommended within the realm of credit risk models to correctly gauge the effect of the economy (global and local) on the credit worthiness of the portfolios. For instance, yearly change of HPI Levels is commonly used in models instead of levels and such transformation is not significantly different from approaches practiced in similar countries in the region.

Are data transformations based on business intuition?

Yes it is based on business intuition. If the splines have different trends, we only keep the one that aligns with business intuition. A good example to illustrate here would be the monotonic ranking of borrower marked-to-market LTV with a floor set at 100 with respect to the LGD model (log odds ratio). For additional details, please see Section 3.1.

Do data transformations satisfy statistical requirements?

[If appropriate, describe statistical requirements for data transformation.]

The statistical criteria for selecting transformed variables are the same as other variables 1) correct sign 2) significance 3) improved model performance and most importantly 4) align with common business intuition. If any of the transformed variables is insignificant ($p>0.05$), has small chi-square or has high correlation (VIF> 50%) with other key predictors, it will be dropped from the model.

Are data transformations applied consistently across similar segments, particularly in the case of macroeconomic variables?

Example: *The same types of data transformations were applied across all segments.*

Yes the variables input pool for both models (PD + LGD) are defined consistently across all segments.

4.1.7 Data sufficiency

What data quality checks have been done?

[Provide evidence of consistency and integrity checks and describe how the data was tested. The tests should include final development data, as well as account-level data.

Data should be analyzed for sudden changes, missing and outlier values, and inconsistent fields within and across records.

Data quality checks should address the following areas:

1. **Business data quality checks –** quick overview of data based on charts to visually detect any outliers or anomalies
2. **Technical data quality checks** – detailed data analysis including:
    - Primary keys and duplicates check
    - Statistical checks -- univariate distribution (mean, median, minimum, maximum, number of observations, and where applicable frequencies) for segmentation, dependent, final independent variables
    - Missing or zero values check, date / time conformity, value format conformity
    - Account level data checks (checks for default values, duplicates, record mapping -- joins, completeness of fields, data and business logic integrity within and across records and frequency checks)

Data quality checks are specified in the following templates:

Data quality checks must include interpretation and analysis of results. Reconciliation is not a data quality check.]

CAMU works closely with the MEP Optima team to ensure data integrity. CAMU examine the distribution of data for both continuous and binary variables, identify outliers and anomalies, and provide treatment to improve data quality. There are two types of data quality checks performed namely: Business Data Quality Check and Technical Data Quality Check.

The Business Data Quality Check provides an overview of the portfolio including key attributes over the performance period consistent in the development time frame of the model. In addition, key portfolio segmentation was considered to identify the robustness and data quality over the same time period. Upon examining the portfolio trend, these key attributes and segments are observed with potential anomalies and portfolio breaks, providing sufficient justification and explanation of potential inconsistency of the data. Data

Quality manager and Business owner or Model Sponsor, attest and review the overall trend to make sure that all observations are consistent with the observed portfolio trend.

The other data quality check performed is the Technical Data Quality, this process consist of examining the development data and the original source data, including original data attributes used in the model development, primary keys, statistical checking and account level check with the data source subject matter experts (SME). Variables used in the development sample are verified and went through a vetting process, checking for inconsistency and data issues. This process includes working closely with the MEP Optima team, to verify frequency check, data consistency, completeness, logic and conformity. Once the account level information has been tested and DM team records the frequency counts and matching for all the variables and explains the results and further attest for the quality of the data for further use in model development. Upon testing, DM team provides feedback to the modelers and data quality managers to address issues or confirm reconciliation results.

For further detail results, please see attachment files named '4.1.7 CCAR Business Data Quality…' for Business Data Quality and '4.1.7 CCAR 2018 Technical Data Quality - Optima VarLib - Severity_CMI_pt1_pt2' for Technical Data Quality results.

What statistical tests were run to verify whether the available data quality is sufficient to develop the model? What were the results of those tests?

[Where applicable, refer to Gating Principles.]

Various Statistical Tests are conduct by the modelers to identify outliers and completeness of the data. Frequency testing, Mean and Standard Deviation analysis are performed as well as testing outlined in the Gating Principles (see Chapter 3.2.1).

Were there any significant events that have affected data quality in the past?

[Describe any known events that might affect data quality. These should not be economic or common portfolio changes but rather technical events such as data lost during data migration.]

Not applicable. The data used for model development are sourced from official and approved data sources with appropriate formal control and data quality check in place. CAMU data quality and Risk DataMart team performed detail business and technical data quality check as part of the maker-checker control process and data variance remained within established control threshold. Other observed data issues are discussed and documented in more detail in the Data Assumption and Data Limitation section in 4.2 and the corresponding mitigation applied.

Is the data coverage sufficient?

[The modeler should engage businesses to understand whether data coverage is sufficient. In cases where the data sample is limited, the modeler must determine whether available external data can augment Bank's internal data.

Aggregated industry data is one option for augmenting Bank's internal data. In order to use industry data, the modeler must first characterize the data from a business and statistical perspective, in order to validate that it is appropriate for the intended use. If the length of Bank's internal data source is too short in duration and if the modeler is unable to source external data that is comparable to Bank's internal data in both business function and statistical properties, then the modeler should proceed to document data shortcomings, including sign-offs from the business that additional relevant data is not available. If a model is developed for areas with insufficient data, please justify and explain.]

Yes. Overall development data performance window has been extended to a cover both pre-crisis and post crisis performance which capture sufficiently full business cycle performance during the stress and recovery period.

For NA Mortgage Method A Residential Mortgage model, observation and performance window used in the development extend from Feb 2006 to Dec 2017 and Jan 2009 to Jun 2017, for the Severity model covering both pre-crisis environment and recovery period post 2008. This rich data coverage captured all the policy changes, business strategy and customer behavior within the full business cycle of the mortgage portfolio.

4.1.8 Vendor-model specific inputs

[If a vendor model is used, please provide the tuning and dialing parameters.

This should include parameters and settings of vendor software needed to replicate loading data into the vendor model such as data filters (within the vendor model), exclusions of certain data ranges and so on.]

The severity model is not a vendor model. There are no vendor-model specific inputs.

4.1.9 Code and Data Guidance

[Model Sponsors must adhere to the coding practices and standards as mentioned in the Code and Data section of the Model Testing Guidance. The information regarding the flow of execution of codes and datasets must be provided in the attached template. All the materials must be shared as per the data sharing process. All the embedded templates should be mandatorily provided. Please refer to code and data guidance section in the Model Testing Guidance for more details]

- Provide details on what software was used for model development.

The SAS programming language was used in model development.

- Provide the location and access to all the codes and datasets. Also, attach the code run book as required by the code and data guidance.

Relevant code and datasets were uploaded to sub-directories within the following MRM server location /ccr/ccar_mvg/1_businessdata/US_Secured/ccar2019/NA_Mtg_Method_A/. Please refer to attachments named '4.1.9….xlsx' for detailed descriptions of uploaded contents.

Model_Testing_Information.xlsx

4.2 Data Assumptions and Data Limitations

- What are the key performance, segmentation and account data assumptions?

  [Include any implied data assumptions or expectations regarding data use.]

The implicit understanding for the model development data is that the loan-level, macro-economic and bureau information (if applicable) used for segmentation and performance measurement (where applicable) is accurate and reflects true account statuses and other characteristics that affect credit performances across time.

It is important to note here that there has been a significant improvement in data-quality control and checks, since the transition to Optima environment and as such the data has been deemed to be robust enough to support CCAR requirements.

As such, CAMU had compiled all the relevant data sources together in one comprehensive document, with description of the data type, type(s) of changes made to the data, frequency of such changes, & archival policy details. This compiled data had been validated by CAMU's internal Data Quality team and subsequently transferred to a single data repository which had been time-stamped and frozen to minimize data pull errors and inconsistencies associated with periodic data refresh. All of these data quality management initiatives that CAMU conducted helped foster consistency and comparability across independent model suites and model usages and aided the built of a model suite that would be able to capture the portfolio trends and the associated risks.

Table 4.2.1: Data Type Overview

#	Data Type	Type(s) of changes over time	Frequency of changes	Archival frequency ?	Examples
1	Static Loan Level Data	1. Historical Restatements	1. Rare restatements	Monthly	Loan Origination Channel: Correspondent, Retail, Wall Street etc.

2	Dynamic Loan Level Data	**1. Incrementally added records** 2. Historical Restatements	1. Monthly 2. Rare restatements	Monthly	Bureau Data, Delinquency status: bk1, bk2 etc.
3	Cumulative Loan Level Data	**1. Event based** 2. Historical Restatements	1. Event based 2. Rare restatements	Monthly	INACTIVE DETAIL: Charge Off, Pay Off etc.
4	Home Price Index	**1. Incrementally added records** 2. Historical Restatements	1. Monthly 2. Frequent restatements	Monthly	N/A
5	Unemployment Rate	**1. Incrementally added records** 2. Historical Restatements	1. Monthly 2. Rare restatements	Monthly	N/A
6	Interest Rate	**1. Incrementally added records** 2. Historical Restatements	1. Semi-annually/Quarterly 2. Rare restatements	Monthly	N/A
7	GDP	**1. Incrementally added records** 2. Historical Restatements	1. Semi-annually/Quarterly 2. Rare restatements	Monthly	N/A
8	Income	**1. Incrementally added records** 2. Historical Restatements	1. Semi-annually/Quarterly 2. Rare restatements	Monthly	N/A
9	SP 500 Index	**1. Incrementally added records** 2. Historical Restatements	1. Semi-annually/Quarterly 2. Rare restatements	Monthly	N/A
10	Volatility Index	**1. Incrementally added records** 2. Historical Restatements	1. Semi-annually/Quarterly 2. Rare restatements	Monthly	N/A

Please note that all the files have been internally verified and frozen as of the last modified date to negate the possibilities of data update related errors/inconsistencies.

Other key assumptions which relate to modeling inputs are presented below.

1. The data used for the CCAR modeling exercise is considered sufficient to fulfil all CCAR and Non-CCAR reporting requirements.

2. The business's current portfolio composition and volume are reflective of the go-forward states.

3. The model is dependent on future macro-economic values. Even if it were perfect, the accuracy of the model is only as good as that of macro-economic forecasts used to operationalize it.

- What are the key macroeconomic data assumptions?

  [Include any implied data assumptions or expectations regarding data use.]

The key macro-economic data assumption is that they reflect the true underlying economic conditions. The Method A RM model suite leverages several macroeconomic attributes to gauge the effect of these economic indicators on the credit worthiness of the real estate portfolio. In the case of interest rates, which are market determined and recorded regularly, this constitutes a small modeling assumption. The importance of the unemployment rate assumption which is official government data based on surveys is comparatively much larger. Similar to unemployment level, state level income is based on the information gathered by the respective states from their Current Population Annual Survey which is based on voluntary participation of the state residents. The assumption effect is largest for HPI information which is proprietary and vendor supplied (CoreLogic - https://www.corelogic.com/downloadable-docs/solutions/loan-performance-secondary-market-analytics-for-capital-markets/capital-markets-real-estate-analytics-suite.pdf, - https://www.corelogic.com/products/corelogic-hpi.aspx). In addition, for both Unemployment and HPI, there is an additional implicit assumption that the appropriate geographic level granularity is optimal. Unemployment, GDP and Income are used at the state level, and HPI, is measured at the CBSA level, which consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people along with all adjacent counties that are socioeconomically tied to the urban center by commuting. The S&P500 and VIX (tested but not used in the final model specification) are measured at the US level.

- What are the key limitations of the data?

  [Include any data limitations.]

The implicit understanding for the model development data is that the loan-level details used for performance, risk drivers and segmentation (where applicable) are accurate and reflects true account status and characteristics across time.

It is important to note here, that with the migration of data to the SAS Grid/Optima/MEP environment, there has been a significant improvement in data-quality control and checks. The Optima environment leverages a centralized data source which undergoes robust data quality checks that adheres to a strict data requirement design which is mandatory for all CCAR models. This has been evidenced in detail in the data quality principles (see 4.1.1 Data Sources and Controls section for pertinent details).

CAMU acknowledges the importance of having a holistic view of the data used in the model development scope to build a robust model. The accuracy of data and transparency of related controls and quality checks around it drive the choice of models built, the performance metrics generated by the system as well as the usability of the end-to-end system by the model users. Therefore, a good ontology of the data is extremely critical for the quality of the work.

To determine if there exist true limitations around the model development data, CAMU diligently researched and executed the following checks to make sure that the data used for the 2019 CCAR process adhered to the highest quality. These checks are listed as below-

- Verified the authenticity and topography of all the listed data attributes.
- Assessed the scope of the data, especially over time, so that the model could avoid the seasonality bias.
- Checked for missing values, identify them, and assess their impact on the overall analysis
- Confirmed that the available pool of development data was large enough to build a robust model. For portfolios or segments that demonstrated sparse data, proactive quantitative decisions were made to use special analytical framework (coping mechanism) to model these loans.
- Made sure data type (numeric, character variables and so forth) is correct and set the upper and lower bounds of possible values.
- Paid extra attention to data integration given that the data comes from multiple sources (REL Datamart, GCRM Office for macroeconomic inputs, CoreLogic for HPI, etc.)

Apart from the quality checks listed above, CAMU did additional due diligence on its own or by collaborating with relevant business partners/teams to understand the reason behind missing values, inconsistencies in the data, presence of duplicates or outliers. Based on further discussions with the business partners, model end users and model reviewers, mitigating actions have been developed to address these data discrepancies to support the CCAR modeling requirements. As such steps were taken in conjunction with the model end users and data subject matter expects (SME) to ensure that the final data set was robust enough for model development. This was either achieved by normalizing/transforming the existing data to suit the purpose of model development and/or creating/leveraging a consistent set of logic to define populations segments across all models and its usages. All these steps were appropriately documented and approvals obtained before model is put into use.

Some of the prominent data weaknesses noticed this CCAR cycle are listed below, along with the corresponding mitigation plans that was adopted-

1. **Low quality of trial information for certain modification program:** The trial information such as month in trial, success/fail in trial, payment and term etc., are the proven to be valuable risk drivers especially for loans in the middle/high delinquency buckets. The predicting power for

such attributes may be reduced somewhat because of the trial information is noisy or sometimes even completely miss-recorded.

Mitigation: Trial information data has improved overtime. Furthermore, few mods are expected in the future.

2. **Missing bureau information:** Significantly higher percentage of missing values for the no-FICO bureau attributes such as "number of open mortgage" (N_M_n_open_mort), "age of oldest trade line" (N_M_age_oldest_trd) as late as April 2009 were observed

   Mitigation: Missing data in the older vintage is not a current phenomenon, which can be addressed in estimation via missing value imputation. The model development data had more than adequate coverage in terms of sufficient parameters and other variables to capture the credit risk of the portfolio. Sufficient care has been taken to minimize the use of the bureau attributes by assessing their overall impact to model performance. Bureau attributes that are not adding much to the model's performance have been dropped. For bureau attributes that are considered important and has missing values, these missing values have been imputed using a median value, thus reducing the potential of omission bias due to omitting them.

3. **Resets around revolving products (ARM/IO)**: Given the rising rate environment, there is an imminent chance of reset risks for specific home loan programs, such as adjustable-rate mortgages or interest-only mortgages that start with a low initial teaser rate and have the potential to reset to much higher rates once the fixed period ends. Such payment shocks are considered a risk factor within the model.

   Mitigation: The model leverages several reset/payment shock attributes to assess the riskiness associated with reset. Specific timing and magnitude variables have been created which measures and quantifies the effect of resets on the loan's performance pre and post (within six months of payment shock) reset.

4. **Recognition of One Timer Loss Events:** One timer loss events are considered outlier observation(s) since they do not represent a normal business process. Initial data exploratory analysis on the loss data had revealed some unusual losses which resulted into significant loss amount or corresponding loss adjustment in a given time period. This was verified by the Risk Portfolio team, the Model Sponsor, who recognized similar events in their loss reporting structure. The following reasons have been narrowed down as the probable causation(s) for these one-time events.

   A) Adjustment due to principal write-downs,
   B) Modification Impact
   C) HELOC conversion,
   D) Credit from legal settlements,
   E) FFIEC Adjustments(Contra-accounting)
   F) Gain on sales

Mitigation: While relevant for the LGD model, one-timer loss events were imputed and removed from the development sample since they do not represent the normal business process. Initial analysis was performed and verified by model user in identifying the one timer list. For additional information on one-timer events and treatments, please refer to attachment "3.2_Memorandum_One-Timer Exclusions - 111417.docx".

5. **Exclusion of data** - The exclusion of data from 2012/01-2014/03 in the PD estimation sample. The exclusion of data prior to 2008 in the LGD estimation sample. The exclusion of data prior to 2009 in the DV Logic Haircut assumption estimation.
   Mitigation - MRM cited a limitation on the current model for not using most recent data to capture recent portfolio mix, performance behavior and macroeconomic trend with the exclusion of the data from 2012/01-2014/03 in the PD estimation sample, data prior to 2008 in the LGD estimation sample and the exclusion of data prior to 2009 in the Haircut assumption estimation. CAMU agrees with MRM that the recent data should be included in the model development sample. The quality of the data prior to 2008 is weak, 2008-2010 is CMI portfolio stress period and needed for a robust stress model. These aspects limit the use of two ends of development data for OOT exclusions. That OOT validation is a regulatory mandate. Therefore, the choice was always limited to pick up a phase in between. The period selected for OOT is excluded from development while variable stability analysis was performed on the full data and various back tests were performed to show overall satisfactory performance across time. Please refer to Section 4.1.3 for additional justification around the OOT data sample.

- Were any exogenous events/shocks evident in the historical data? How were they accounted for?

  Exogenous events such as one-time accounting loss restatements have no effect on the development of the transition model which is built on delinquency and terminal events. These one-time losses are excluded from the LGD model development to reduce biases within the model's performance. Please see LGD Development data waterfall in MDD Section 4.1.3. CAMU's modeling practices align with Bank's accounting policies as reflected in Global Consumer Credit Fraud Risk Policy (GCCFRP) Chapter 20.

- What factors, if any, mitigate the limitations of the data?

  [Include a sub-section for each identified limitation of the data (e.g., short data period, potential inconsistencies, etc.). If none exists, provide a rationale for why one is not needed.

  Please see mitigation plans attached as part of the data limitations discussion, as enumerated above.

5. Model Specification

[Include an explanation of the theory, logic, and design behind each model, [and] a description of model selection and specification, variable choice, and estimation methodology.]

The goal of this chapter is to provide a clear description of the modeling framework and final model specification(s). The chapter includes a description of how the framework was derived, as well as an overview of assumptions and limitations that may reduce the accuracy and effectiveness of the model.

This chapter provides a clear description of the modeling framework and final model specification. Included is a description of how the framework was derived, as well as an overview of assumptions and limitations that may reduce the accuracy and effectiveness of the model. First, the model's intended purpose and objective, in technical terms, is provided. Next, the underlying logic behind the selected model is discussed. An analysis of the model output, including the congruence of inputs with the assumed economic scenario, the justification of any qualitative adjustment, along with the statistical analysis used to support the model output is then documented. Further, the rationale for portfolio segmentation and variable selection is included. In addition, final model specification is discussed along with model specific assumptions and limitations.

5.1 Model Methodology Overview

The purpose of this Method A credit loss model is to estimate Gross and Net Credit Losses (GCL and NCL) – units, dollar and as percent of ending net receivables', intermediate delinquency buckets (units and dollars) projections and Voluntary Payoff projections (units and dollars) for the North America Residential Mortgage portfolio. The type and nature of the reportable outputs vary across the model's stipulated usages. The forecast horizon and the scenarios modeled also vary for specific outputs given its stated usages. For more information on the model's usages, please refer to the Model Usage Grid.

5.1.1 Model Objective

[Describe the model's intended purpose and objectives in technical terms and in the context of the regulatory requirement.]

This model aims to project monthly Gross and Net Credit Losses for the modeled portfolios over a given time horizon conditional on macroeconomic scenario assumptions over that horizon. Apart from the credit losses, the model is also equipped to provide delinquency and voluntary payoff forecasts for the stipulated model usages. The overall modeling approach follows an account level, competing risk framework where

Expected \$ Net Credit Loss = Probability of Default × \$Exposure at Default x Loss Given Default, or

$$\$NCL = PD \times \$EAD \times LGD$$

The underlying section provides a detailed description of the design, theory and logic surrounding this modeling framework.

5.1.2 Design, Theory, and Logic of the Model

[Provide the underlying logic behind the selected model. If the modeling approach is novel, take special care to clearly and comprehensively discuss the motivation, key decision making, and developer thought processes behind selecting the approach.

The description should be detailed. Articulate the modeling methodology and model processing components, and include mathematical specification and numerical techniques and the approximations employed in the model design. In addition to describing the overall modeling approach, directly address by explaining and justifying the approaches selected to handle specific modeling issues or concerns, for example, the modeling approach used to address loan servicing effects such as modifications, and the modeling approach used for bankruptcy losses/defaults, whether or not bankruptcy losses/defaults were modeled separately from other types of losses/defaults.]

The 2019 Method A Residential Mortgage PD Transition Model is a loan-level model designed to produce forecasts for delinquency status as well as both voluntary and involuntary termination.

The overall modeling framework is set in a competing risk framework that calculates the monthly transition probabilities over the target horizon as a function of both loan-specific characteristics and macroeconomic factors. Given these probabilities, the model framework then assigns a loan level expected loss given default (LGD Rate) and exposure at default (EAD) as summarized below. Finally, this loan level expected losses are aggregated across all loans to produce an estimate of end-to-end portfolio-level losses. The overall framework can be illustrated using the specific mathematical function-

$$EL_{it} = PD_{it} \text{ (Prob. Of default)} \times LGD_{it} \text{ (loss given default)} \times EAD_{it} \text{ (Exposure at default)}.$$

where ***t*** represents time period and ***i*** represents loans.

Probability of Default (PD) component

CAMU used a transition model for estimating the probability of loan default. A transition model is appropriate for situations when an account moves through different delinquency states each month. For example, an account in a “current" state this month will be in a “current" state next month if a payment is made by the due date, and will be in a “30 days past due" state if no payment is received. Another valuable feature of this design is that the transition model maintains the progression and timing of events in the path from “current" to “loss.

The path independence feature of the transition chain model is effectively relaxed in the transition model framework by including delinquency history variables in the regression, and allowing future transitional probabilities to be affected by the prior months’ values.

The transition model represents the month-by-month movement of loans between current, delinquency, VP, and IVP states. The risk horizon is one month; therefore, the range of values that the loan can take at the end of that period is characterized. Let us first list all possible credit outcomes that can occur at the end of the month due to credit events

- The loan stays at current state
- The loan migrates to other nonterminal state; or
- The loan moves to VP or IVP terminal state.

Each outcome above has a different likelihood or probability of occurring.

CAMU's modeling framework proposes that the multinomial probability for any row of the transition matrix can be estimated by a collection of binary logistic models using the results from Begg and Gray *(1984 Biometrika; 71:11-18)*. CAMU estimated binary logistic regression models of transition relative to a reference delinquency status as a function of independent variables or co-variates. The dependent variable is the migration or transition rate from a delinquency bucket at the observation month to a delinquency bucket of choice rather than to the reference delinquency bucket. Collateral attributes, payment pattern, borrower credit characteristics, loan features and macroeconomic factors - changes in home prices, unemployment rates and interest rates - influence the loan-level transition rates. There are two terminal states for a loan: Voluntary Payoff (prepayment and contractual termination) and Involuntary Payoff, IVP (also known as default).

One can demonstrate the model with few mathematical constructs.

As in other forms of linear regression, multinomial logistic regression uses a linear predictor function $f(j,k)$ to predict the probability that observation j has outcome k, of the following form:

$$f(j,k) = \beta_{j,k,0} + \beta_{j,k,1}x_{j,k,1} + \beta_{j,k,2}x_{j,k,2} + \cdots + \beta_{j,k,m}x_{j,k,m} + \cdots + \beta_{j,k,M_{j,k}}x_{j,k,M_{j,k}}$$

Where $\beta_{j,k,m}$ is a regression coefficient associated with the mth explanatory variable for moving from state j to the kth outcome, $M_{j,k}$ is the number of explanatory variables for this j to k move.

As explained in Begg and Gray's logistic regression work, the regression coefficients and explanatory variables are normally grouped into vectors of size $M_{j,k} + 1$, so that the predictor function can be written more compactly:

$$f(j,k) = \beta_{j,k} \cdot X_{j,k}$$

Where,

$\beta_{j,k} = \left[\beta_{j,k,0},\ \beta_{j,k,1},\ \beta_{j,k,2}, \cdots,\ \beta_{j,k,m}, \cdots,\ \beta_{j,k,M_{j,k}}\right]$ is the set of regression coefficients associated with moving from j to k, and

$X_{j,k} = \left[1,\ x_{j,k,1},\ x_{j,k,1},\ x_{j,k,2}, \cdots, x_{j,k,m}, \cdots, x_{j,k,M_{j,k}}\right]$ (a row vector) is the set of explanatory variables associated with move from state j to k.

Let p(j, k) for k = 1, . . . , K-1 denote the probabilities of moving from the jth state this month to the kth state next month. And p(j,K) denote the probabilities of moving from the jth state this month to the Kth (reference) state next month. Writing the multinomial logistic for the jth row as separate binary logistic models yields the collection

$$\frac{P[\{\text{state k next month}\}|\{\text{state j this month}\}]}{P[\{\text{state K next month}\}|\{\text{state j this month}\}]} = \exp\left(\beta_{j,k} \cdot X'_{j,k}\ \right), k = 1, 2, \ldots, K-1.$$

Where,

$X'_{j,k} = \left[1,\ x'_{j,k,1},\ x'_{j,k,1},\ x'_{j,k,2}, \cdots, x'_{j,k,m}, \cdots, x'_{j,k,M_{j,k}}\right]$ are inputs to the model. K is the reference state.

To introduce a loan-level model for the transition probability $p(j,k)$ with explanatory variables $X_{j,k} = \left[1,\ x_{j,k,1},\ x_{j,k,2},\ \cdots, x_{j,k,m}, \cdots, x_{j,k,M_{j,k}}\right]$, k= $1, 2, \ldots, K-1$, use a sample of loans that are in state j this month and move next month to either state k (Y = 1) or state K (Y = 0). The logistic model estimated from this sample is denoted by $\hat{\beta}_{j,k}.X'_{j,k}$, Where $X'_{j,k}$ is the model inputs for $X_{j,k}$, $\hat{\beta}_{j,k}$ is the estimate for $\beta_{j,k}$.

It's important to stress that this is not the transition probability and more is required to determine $p(j,k)$ from the above binary logistic regression prediction. Within the multinomial outcome, using the condition that all k of the probabilities must sum to one, then

$$P[\{\text{state k }\}|(\{\text{state j this month and state k next month}\}\text{or }\{\text{state j this month and state K next month}\}\ )]$$

$$= \frac{\exp\left(\hat{\beta}_{j,k} \cdot X'_{j,k}\right)}{1 + \sum_{k^*=1}^{K-1} \exp\left(\hat{\beta}_{j,k^*} \cdot X'_{j,k^*}\right)}$$

$$P[\{\text{state K}\}|(\{\text{state j this month and state k next month}\}\text{ or }\{\text{state j this month and state K next month}\}\ )]$$

$$= \frac{1}{1 + \sum_{k^*=1}^{K-1} \exp\left(\hat{\beta}_{j,k^*} \cdot X'_{j,k^*}\right)}$$

PD Implementation / Forecast Design: Expected Probability Approach

When a transition matrix model is used to predict the probability of default (PD), there are typically two ways to implement the model: Expected Probability approach and the Simulation approach. The 2019 Method A model suite continues to utilize the Expected Probability Approach for all forecasting needs.

In the expected probability approach, the probability distributions associated with the output variables are calculated from the probability distributions associated with the input variables. In the simulation approach, the value of a distributed parameter is selected by the generation of a random number, with the probability of a given value being determined by the association of random numbers to that variable. By repeating this process a large number of times, a picture of the distribution of the output random variable may be built up, from which estimates of the parameters of interest may be calculated, e.g. their mean, standard deviations, etc.

First and foremost, using the Expected Probability approach, one can easily get the loan level cumulative probabilities for terminal (voluntary and involuntary prepayment) events as well as the probability distribution for active loan statuses for each forecast period. Once the modeling framework is generated, outputs can be rapidly processed with or without the use of a computer. This makes the expected probability approach more user-friendly as it can be used to generate any type of intermediate/terminal outputs without adding any additional run time to the model.

Second, the expected probability approach gives exact model results given the assumptions used in the model. Within the simulation framework, there is virtually no limit to analysis as an infinite number of paths can be generated using the model's simulation feature. It is possible for simulation results to show widely different values for the same model when simulated with different random number seeds. The system itself has to be run several times in order to get reliable parameter estimates.

Third, under the simulation approach it takes multiple simulations for the results to converge. This often increases the operational costs and time associated with forecasting the results. The expected probability approach, given its structural framework, avoids this problem.

Exposure at Default

Exposure at default is designed to calculate performing balances, defaulted balances and expected losses etc. through each month of a scenario horizon. The EAD framework is constructed as a balance flow tree that computes the expected performing and defaulted loan balances through time. All loans are classified as either "amortizing" or "non-amortizing". It is assumed that all amortizing loans will continue to pay down under the amortization schedule which covers both the principal and interest due on the balance at any given time, using the inputs like the term remaining and interest rate on the loan. In the EAD framework, for non-amortizing positions, the amortization deduction is skipped when calculating balances over the forecast horizon. Each month, PD is used to estimate the proportion of the previous month's performing balance that will be in default in the next period, while cumulative partial charge-off to-date compared with total severity at liquidation is used to estimate the proportion of the previous month's defaulted balanced that is expected to reverse. After applying default and cure flows, interest is charged on the performing balances. The performing balance is then amortized according to the amortization rates described above, except for non-amortizing contracts such as interest-only loans. The remaining balance is reduced by an amount of expected prepayments determined by the prepayment rate forecast. Only performing loans incur amortization and prepayments. Non-performing loans are assumed to cure prior to prepayment.

An independent review of the CCAR 2018 Method A Model Suite code package conducted by the Independent Risk Modeling Oversight (IRMO) team in Q1 2018 revealed some discrepancies within the code logic. CAMU conducted research and analysis to confirm accuracy of the observations, identified potential root causes, and assessed impacts to model results from any changes proposed. The items impacting model accuracy have been implemented as part of the 2019 CCAR Method A submission. Please note, that prior to implementing the proposed changes in the CCAR 2019 Method A code package, a memorandum was shared with Risk business partners and model end users for awareness and assessment of the proposed recommendations.

One of the key areas from this assessment was the observations related to amortization calculations and treatment of loan principal balance amounts. The observations that were specific to the residential mortgage portfolio pertained to continued application of an amortization factor to loans in the bk7_bk7 transition and overestimation of amortization for the recovering loans. Given these objective observations and subsequent review assessments, CAMU created and implemented a new amortization module that replaced the existing amortization factor based solution with a cash flow based solution to estimate the unpaid principal balances. This new amortization module holistically address how the handle related to amortization for deep delinquent and recovering loans.

The following enhancements have been introduced with the new module for the 2019 CCAR submission process:

- Amortization will utilize the contractual *payment* of the loan. The contractual payment, as of the portfolio snapshot, is available from the OPTIMA source data. For fixed-rate non-IO loans, this payment is sufficient. For ARM loans or loans beginning their amortization, rate resets and payment calculations can be forecasted based upon contractual terms.
- Loan balances evolve according to the following algorithm when borrowers make their contractual payment.

$$Interest_t = UPB_{t-1} * r_t/1200$$

$$Principal_t = \begin{cases} UPB_{t-1} & t = Termination \\ min(UPB_{t-1}, Payment_t - Interest_t) & otherwise \end{cases}$$

$$UPB_t = UPB_{t-1} - Principal_t$$

The principal payment is the total principal and interest payment (available from source data or forecasted for ARM loans) less the accrued interest, but capped by the outstanding balance. Furthermore, multiple payments (e.g., curing of a delinquent loan) may be achieved through multiple iterations of the above algorithm.

Beyond simple amortization of the loan assuming that it makes scheduled payments, Method A requires expected balances for loans when in the various delinquency statuses. Let the (expected) balance of a loan that is in state $s \in \mathcal{F} = \{0,1,2,3,4,5,6,7\}$ at period t be denoted as $bal_{s,t}$. The current production code computes $bal_{s,t}$ recursively, as follows:

$$bal_{s,t} = \sum_{k \in \mathcal{F}} T_t^{k \to s} bal_{k,t-1} F^{\max(0,k-s+1)}$$

where $T_t^{k \to s}$ represents the (conditional) probability of transitioning from state k to state s in period t. Letting s_t denote the state of a loan at period t, we can restate the definition according to the following:

$$T_t^{k \to s} \equiv P(\, s_t = s |\, s_{t-1} = k)$$

F represents a one-period amortization factor; $k - s + 1$ denotes the number of payments needed to transition from state k to state s over a single period.

CAMU has also changed the balance calculation to the following:

$$bal_{s,t} = \; P(\, s_t = s) * UPB_{s,t}$$

where $UPB_{s,t}$ is the (deterministic) balance of a loan in state s at period t. This approach is simpler and preferable for the following reasons:

- It requires only the unconditional state probabilities $P(\, s_t = s)$ and deterministic balances.
- The balances can be calculated by amortizing a current (i.e., non-delinquent) loan and keeping track of lagged values. This is due to the following relationship:

  $$UPB_{s,t} = UPB_{s-1,t-1}$$

  Therefore, the (deterministic) balance of a loan at period t in the various states can be determined purely through lagged values of a current loan.

$$\begin{aligned}
UPB_{1,t} &= UPB_{0,t-1} \\
UPB_{2,t} &= UPB_{0,t-2} \\
UPB_{3,t} &= UPB_{0,t-3} \\
UPB_{4,t} &= UPB_{0,t-4} \\
UPB_{5,t} &= UPB_{0,t-5} \\
UPB_{6,t} &= UPB_{0,t-6} \\
UPB_{7,t} &= UPB_{0,t-7}
\end{aligned}$$

- Balloon loans are implicitly handled. Since balloon loans are typically fixed rate, usage of the contractual payment, as calculated by the servicing system and fed through OPTIMA, effectively makes the distinction between the amortization and the maturity terms.
- Eliminating the amortization payment for bucket 7 to bucket 7 transitions is easily accomplished by adding back this payment.

$$bal_{7,t} = P(s_t = 7) * UPB_{7,t} + P(s_{t-1} = 7) * T_t^{7\to 7} * \left(UPB_{7,t-1} - UPB_{7,t}\right)$$

Please refer to the pseudo loan amortization example within the attachment- '1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY.docx' to illustrate the set of algorithm above.

Distressed Valuation Haircut Lookup Logic (DV Logic)

In real estate, a property that's in the process of foreclosure is generally referred to as being distressed. Owners of foreclosure properties may consider purchase offers well below what they'd normally accept under normal circumstances. In the case of real property in foreclosure, its distress value may be much lower than its current true market, appraised and tax-assessed values. For the 2019 CCAR modeling process the Method A model suite has utilized completely redeveloped distressed valuation logic to apply mark-to-market property value haircut on distressed properties based on home price movements in conjunction with loan level attributes.

To provide some background, the Method A model suite in prior years had utilized a fully functional DVM model to estimate future distressed property valuation(s) (DVs) for:

- Loans with an existing distressed valuation on the underlying collateral (referred as the PDV=prior distressed value segment),
- Loans with no existing distressed valuation (referred to as the NPDV = no prior distressed value segment), at the time of the forecast

This DVM model was used in combination with loan-level PD transition models to predict the losses for both CCAR and Non-CCAR usages of the model. This model covered all of the NA Mortgage (Residential and Home Equity) which included the three lines of businesses – BankMortgage (CMI), Bank Private Bank (CPB) and US Retail Bank (USRB). The DV model used in prior year's forecasting processes served two main purposes:

1. The model was to calculate the mark-to-market Loan to Value (LTV) on the property, which is a key attribute for the Severity (LGD – Loss Given Default) model, and,
2. The month-over-month DV deterioration is used to calculate incremental pre-liquidation FFIEC write-downs (also referred to as Partial Charge Offs – PCO) after the loan is predicted to go 180+ days past due

With a simple design to apply loan level collateral valuation updates from home price shifts, the prior year's DVM model had demonstrated reasonable back testing performance and sufficient sensitivity in stress scenario forecasts over the past three years. However, the Independent Risk Modeling Oversight

(IRMO) team had shared functional review observations that the entire modeling framework was hinged on a single macro-economic attribute – HPI and a time decay factor. Conclusively, as part of the Model Validation Report (MVRT), it was recommended to enhance the variable selection process to consider business and loan level attributes. Apart from the MVRT recommendation, the Model Risk Management's (MRM) policy that stipulates a model's aging guidelines (models should be re-estimated every three years or must be granted an exception for continued use), had already led CAMU to prudently reevaluate other alternative modeling approaches for estimating DVs on properties.

At the start of the 2019 model redevelopment process, CAMU discovered several avenues aimed at streamlining data sources as utilized by Collateral Risk Management processes, some of which have been part of the ongoing Mortgage Transformation initiative, as part of Bank's broader strategy to align the mortgage business with retail banking. Given this strategic business initiative which led to an overhaul in systems & processes, CAMU proactively consulted with the Collateral Risk Management and Risk Data Management teams, to identify data sources which could be leveraged for updated property values. Considerable time and efforts were also spent to understand the logistics around the new data files and also assure that these data sources align with Collateral Risk Management's valuation practices. The Attachment '5.1.2 Property Valuation Waterfall Logic' contains details on the Portfolio Valuation Waterfall procedure used by Collateral Risk Management. Distressed Valuation haircut (DV) rates are estimated using the same dataset as what is used for current Collateral Risk Management processes (starting from 2010) as well as observations from a legacy data source (DRI, for observations dated 2009 and earlier).

For the 2019 Method A CCAR process, CAMU proposed a mark-to-market haircut lookup logic that would be based on loan characteristics. The rationale justifying the segmentation of loans into prior distressed valuation (PDV) without prior distressed valuation (NPDV) is based on data availability and commonly perceived intuition. For all properties that had a prior distressed valuation, it is recommended to leverage this value for accurate assessment of the property. However since not all properties have a prior distressed value available, the workaround is to impute a value based on the original home value. The haircut rates have been estimated based on empirical data for both PDV (prior distressed valuation available) and the NPDV (no prior distressed valuation available) and this change has been implemented in conjunction with a PD/LGD model framework revision to directly incorporate home price updates at various stages of the forecasting process. To examine the incremental effect of loan characteristics on the distressed property's value, CAMU created a comprehensive list of loan attributes listed as below, which were tested on a stand-alone basis to evaluate its impact on the DV of the property:

- Lien Position
- Occupancy Type
- Origination Channel
- Property Type
- Loan Balance
- LTV

- Loan Age
- Loan Type (Government versus Conventional)
- PMI Coverage
- Modification Status
- Business Entity (CMI or other)
- State of the property (judicial vs non-judicial)

The attachment –'5.1.2 pivot_NPDV_Haircut_newData_HPIV4byQtr.xlsm' shows the haircut ratio across time for the NPDV segment. It shows among all the dimensions that were tested, only occupancy type and lien position came out significant. Haircut ratio is significantly lower on the owner-occupied property than the non-owner occupied property, and it's also lower on the 2nd lien when compared on first lien.

The attachment –'5.1.2 pivot_PDV_Haircut_newData_HPIV4byQtr.xlsm' shows the haircut ratio across time for the PDV segment. It shows no significant difference in all of the dimensions that were tested. Lastly, the attachment –'5.1.2 DRI_vs_DM_On_PDV.xlsb' presents the side-by-side comparison of the haircut ratio as derived on DRI and DM. It shows that the haircut ratio based on DM data is close to DRI data when both datasets are available. So although the DM data is not available for the stress period, we borrowed the haircut ratio as obtained from the DRI data. Further, as per the DRI data, no significant differences were observed in the haircut ratios across the risk drivers.

This analytical testing, not only aided in understanding and discovering the attributes which had the most impact on the property's value across business cycles and also helped assure the model's stability and usability across all important population sub-segments, key risk drivers and across all backtesting snapshots.

Once the key loan level drivers were identified, an all –inclusive analysis was conducted which further examined the sensitivity of the DV rates to the corresponding changes in the HPI. The sensitivity testing, that were executed were based on the following mathematical annotations for the NPDV and PDV segments respectively-

NPDV Segment:

$$DV(t) = \frac{HPI(t)}{HPI(Orig)} * Orig\_Prop\_Amt * (1 - haircut)$$

PDV Segment:

$$DV(t) = \frac{HPI(t)}{HPI(time\ of\ prior\ DV)} * Amt\ of\ last\ DV * (1 - haircut)$$

Based on the analyzed results, CAMU proposed the following haircut for the following segments-

1. For the PDV segment, a 10% haircut rate is recommended for stress period and no haircut is needed for non-stress period forecasts
2. For NPDV segment, as depicted above, the haircut rate was most sensitive to lien position and occupancy type. Other attributes tested included FICO, LTV, and Property type. first Liens and non-owner occupied properties had a higher haircut rate compared to 2nd liens and owner-occupied properties respectively. The other finding pertaining to this analysis was that the haircut rate for these loans remained fairly stable across stress and non-stress periods which were again based on a combination of both DRI and DM datasets. The following 2x2 matrix illustrates the haircut rates for the NPDV segment across the identified key risk segments.

Table 5.1.2.1 Haircut ratio on NDPV segment

Ratio	first lien	Non-first lien
Owner-Occupied	0.20	0.13
Not Owner-Occupied	0.27	0.22

The business intuition for the selected parameters of the proposed haircuts are as follows –

1. Owner Occupancy – For all properties occupied by owners or those that serve as primary residences, it is always in the best interest of the owners to not subject the property to much depreciation. Properties, which are non-owner occupied, are susceptible to more depreciation, as there is no real equity involved here. Hence, owner occupancy came out to be a significant attribute in the different haircuts chosen.

2. Lien Position – Since first liens usually tend to be the primary residences of the borrowers, it is very typical for the borrowers to default on their second liens first, when they encounter some form of financial hardship. So those defaults on second lien are expected in less stress conditions compared to the first liens.

Please refer to Chapter 3 for a detailed comparison of the Prior DV Model with the new DV Logic. The main alternative to the 2019 DV Logic is the prior year's DV model [Model ID # 108730]. For additional details on the methodology for the 2017 DV model, please refer to attachment –' 5.1.2 Prior DV Modeling Methodology- Alternate Approach'.doc.

Please refer to Chapter 3 for a detailed sensitivity comparison of the Prior DV Model with the new DV Logic.

The Severity Model - Estimating Loss Given Default

When a loan is projected to default (go into IVP state) and the collateral is repossessed, the proceeds from sale of the property are used to offset the balance of the loan and cover the associate expense. In most of the cases the collateral is not sufficient to absorb all associated costs. As a result, a loss may be realized. In this CCAR loss forecasting discussion our focus remained on cumulative accounting credit losses.

The model development team underwent an extensive process to challenge and select the functional form for the 2019 LGD model. In the prior year's model, CAMU leverages a proprietary 2-stage modeling framework for the first lien loans wherein the first stage used a loss disposition selection model in a multinomial framework, while the second stage estimated the loss given default (LGD) for corresponding disposal outcome using non-linear regression framework. However as illustrated within Chapter 3 of the MDD, the loss disposition type is a policy driven mandate which has continued to change and evolve over time as guided by the business's current lending practices, regulatory landscape and the macroeconomic environment. The volatilities associated with the different loss dispositions over time(please refer to Figure 3.1.2: Historical Trend – Loss Disposition Type (first Lien Only) implied that there is no consistent relation between disposition type and actual losses incurred which in turn implies that a historical model fit would not be able to accurately reflect the policy changes over time. Another data specific challenge with modeling the loss disposition type attribute is that the disposition type is not static over time. As illustrated within Chapter 3 (Table 3.1.3: Loss Disposition Type Change over Time) approximately 4.4% of the IVP loans underwent change in disposition type over time. Modeling with a static disposition type, would not lead to correct modeling of disposition types. Again, with the influx of zero loss loans in recent times (influenced by the rising housing prices given a stronger economic outlook), it was considered a modeling fallacy to exclude out the zero loss loans, which the existing framework did.

Given all these considerations (please refer to Chapter 3 for a holistic comparison of the new vs old modeling approaches) and the feedback from model reviewers', the modeling team considered it prudent to replace the existing approach with an approach which has been used for modeling the second lien loans. In this way, the modeling team was able to maintain a consistent modeling framework for both first and second lien loans.

The new modeling framework for modeling both first and second liens consists of two stages: the first stage estimates the probability of losses by their outcome, i.e. the first stage models the probability of incurring full/zero losses relative to partial losses and the second stage estimates the severity for the defaults with partial losses. Full or zero loss loans are loans which experience either full losses equal to the amount of their unpaid principal balances while zero losses are common for loans which had some prior write-downs(FFIEC) and had a reversal of the write-downs due to consistently improving housing prices. Loans with full losses have full losses with a probability of one while zero loss loans experience zero losses with a probability of zero. The partial losses on the other hand experience losses which have a probability between [0, 1]. The first stage model uses linear logistic regression and the maximum likelihood estimation is employed. The second stage model uses nonlinear regression and nonlinear least squares estimation is employed.

As shown in the Figure 5.1.2.1a below, CAMU has implemented a Loss Outcome Based Loss Severity Modeling structure for NA Mortgage first lien and second liens. Under an Outcome Based Loss Severity Model, a two stage modeling process is implemented. In Stage 1, a probabilistic choice is made between full/Zero losses and partial losses. Conditioned on the probability of loss type, a loss rate model is applied to obtain loss severity for the partial loss type loans in the second stage. The final loss estimate is the product of these two stages and balance at default. This is different from the prior year's LGD framework which leveraged a loss choice based modeling framework (Figure 5.1.2.1b), wherein a two stage modeling process was implemented. In Stage 1, a probabilistic choice was made between Real Estate Owned (REO), Short Sale, Third Party Foreclosure, or Charge-Off (loss disposition types) for first liens and between full losses and non-full losses for second liens. Please note zero loss loans were excluded from the model development data in the prior LGD modeling approach. This has been remedied in the 2019 LGD modeling approach, as discussed above and in Section 3.2.1. Conditioned on the probability of loss type, a loss rate model was applied to obtain the loss severity. Loss estimate was calculated as the product of these two stages and balance at default.

Figure 5.1.2.1a: 2019 LGD Framework – Loss Outcome Based Loss Severity Modeling Structure

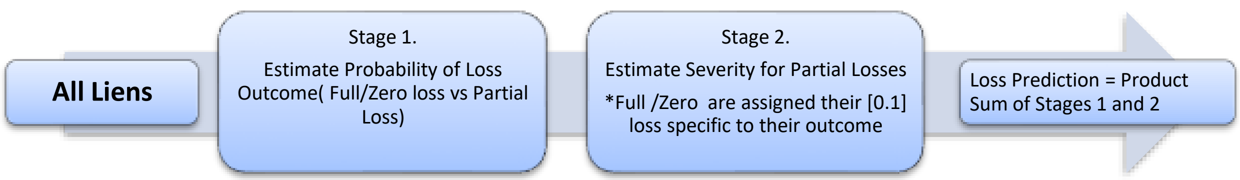
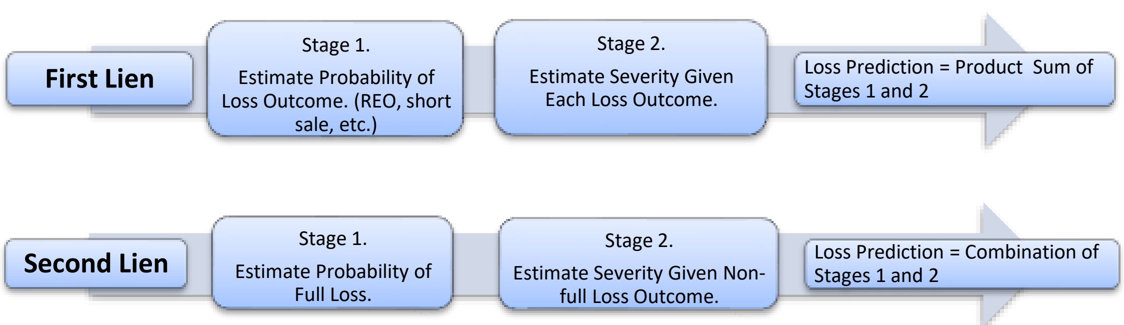


Figure 5.1.2.1b: 2017 LGD Framework –Choice Based Loss Severity Modeling Structure



Under the 2019 LGD modelling framework, the model's sensitivity and monotonicity had witnessed significant improvement with respect to borrower's LTV, considered as one of the key risk drivers of the LGD model. Intuitively speaking, higher LTV is associated with higher severity rate due to relatively higher debt compared to the total property value. Intuitively speaking, periods of rising HPI, indicate improved home valuation which in turn lowers the mark-to-market LTV (current loan balance/home value) and the corresponding loss rates. Under the 2017 modelling framework which modelled all loans by their loss disposition types, the empirical data demonstrated a 'V' shaped relationship between the LTV and loss severity(Please refer to Figure 3.1.4.1 in Section 3.1) which implied that for all first lien loans in the lower LTV bins, as LTV increased, the loss severity decreased. This had a counterintuitive implication given the common business understanding associated with rising HPI. Even after creating a LTV floor at 100, the model continued to render counterintuitive connotation with respect to the LTV variate (see Figure 3.1.4.3). Since a definitive solution could not be devised to nullify the V-shaped effect of borrower's LTV to loss rate, it was deemed necessary to alter the modelling framework for the first lien loans to accurately model the LTV effect for the losses (partial) observed. For additional details on the model's sensitivity and monotonicity to LTV, please refer to Section 3.1 of the MDD.

Presented below is the mathematical representation and logic for the first and second lien loss severity models as development for the 2019 CCAR process.

LGD estimation for non-CPB, non-VA loans

In the first stage model binary logistic regression model are applied to estimate the probability of each loss outcome using partial losses as the reference category. In other words, two binary models are utilized to model the probability of a zero or full loss using partial loss as a reference category. Each logistic model measures the relationship between the dependent variable and a set of key risk drivers (predictors). In each of the models, the dependent variable indicates loans that belong to a certain loss outcome 'j' other than partial (1 = loss type j, where j= 'zero loss' or 'full loss'; 0 = Partial loss).

In the second stage of the model, a nonlinear regression model is then used to model the severity rate conditional on the partial loss. Both zero and Full loss are kept outside of the purview of the second stage model as they usually have a pre-assigned loss rate corresponding to 0 or 1. The partial loss rate on the other hand has a severity between 0 to 1 range and as such the second stage model essentially estimates the β_j or the vector of coefficients in the partial loss model, corresponding to a set of predictors which includes specific loan characteristics and macroeconomic variables.

Stage 1: Two sets of *Binomial Logit Regression to Estimate*

1) Probability of Full Loss vs Partial Loss

$$\boldsymbol{Y_i\,(Full\ Loss) = \begin{Bmatrix} 1\ \ if\ Full\ loss & for\ individual\ i \\ 0\ \ if\ Partial\ Loss & for\ individual\ i \end{Bmatrix}}$$

let $\boldsymbol{\pi_i\,(Full\ Loss) = Prob\{Full\ Loss|\ X_i\} = Prob\{Y_i\,(Full\ Loss) = 1|X_i\}}$

2) Probability of Zero Loss vs Partial Loss

$$\boldsymbol{Y_i\,(Zero\ Loss) = \begin{Bmatrix} 1\ \ if\ Zero\ loss & for\ individual\ i \\ 0\ \ if\ Partial\ Loss & for\ individual\ i \end{Bmatrix}}$$

$$\pi_i\,(Zero\ Loss) = Prob\{Zero\ Loss|\ X_i\} = Prob\{Y_i\,(Zero\ Loss) = 1|X_i\}\ \ ,$$

where X_i is a vector of account level characteristics and other related factors for individual loan i.

For each loss type j we have the simple logistic model form

$$ln\left(\frac{\pi_i}{1-\pi_i}\right) = X_i'\beta$$

where $ln\left(\frac{\pi_i}{1-\pi_i}\right)$ is a logit of the dependent variable and β is a vector of coefficients. The Odds of full loss and zero loss are:

$$Odds(Full\ Loss) = \frac{\pi_i(Full\ Loss)}{1-\pi_i(Full\ Loss)} = e^{X_i'(Full\ Loss)\beta(Full\ Loss)}$$

$$Odds(Zero\ Loss) = \frac{\pi_i(Zero\ loss)}{1-\pi_i(Zero\ Loss)} = e^{X_i'(Zero\ Loss)\beta(Zero\ Loss)}$$

Therefore, probability that a default account i will go through full/zero/partial severity is:

$$Prob(Full\ Loss) = \frac{Odds(Full\ Loss)}{1 + Odds(Full\ Loss) + Odds(Zero\ Loss)}$$

$$Prob(Zero\ Loss) = \frac{Odds(Zero\ Loss)}{1 + Odds(Full\ Loss) + Odds(Zero\ Loss)}$$

$$Prob(Partial\ Loss) = \frac{1}{1 + Odds(Full\ Loss) + Odds(Zero\ loss)}$$

<u>Stage 2</u>: *Non-linear Regressions to Estimate Severity Conditional on partial Loss. Please note both zero and full losses are assigned their respective losses and hence do not need to flow through the second stage model. Severity is defined as cumulative lifetime loss (up to 6-month after IVP date) divided by Balance at IVP.*

Using the partial loss observations, severity is estimated via non-linear regression model

$$Severity_i = \frac{e^{x_i\beta}}{1 + e^{x_i\beta}}$$

where X_i is a vector of predictor variables and β is a vector of associated coefficients.

Combining two stages, we have individual i's loss estimate

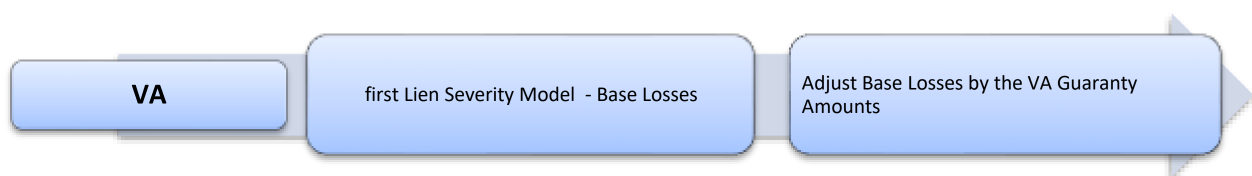
$$\boldsymbol{Loss\ Estimate} = \boldsymbol{Prob}(\boldsymbol{Full\ Loss}) \times \boldsymbol{UPB} + \boldsymbol{Prob}(\boldsymbol{Partial\ Loss}) \times \boldsymbol{Severity} \times \boldsymbol{UPB}$$

VA Model

A VA loan is government-backed, and the complete loss recovery is often possible through VA's full bid, which depends on several factors such as total loan indebtedness and the net value of the property. In case of partial bid, or even no bid for small loans, there still exist chances of full recovery thanks to the coinsurance feature of the guaranty payment schedule. Given the specific criteria for the VA loans, the NCL losses for these loans are estimated based on business rules provided by the model's business users. Specifically, VA loans losses are adjusted based on the specific loan size by the guaranteed amount, as dictated by the relevant accounting policy, maintained by Finance and adhered to by the businesses. This policy is available in form of a lookup table which has been passed on to the model development team for incorporation in the modeling process.

For the 2019 Method A Severity modeling process, it was considered prudent by the model developers to not develop a separate model for VA loans. Hence, while the VA loans were excluded from the model development sample, for all model forecasting purposes, the first lien, severity model was leveraged to forecast the base losses. These base losses was then adjusted by the appropriate guarantee amounts within the production system based on the lookup table which references the allocated guaranteed amount for the given loan size. Please refer to the attachment – '5.1.2 VA Guaranty Amounts' below for the VA Guaranty allocations. Please note that these guarantee amounts are directly coded within the production code as part of model implementation.

Figure 5.1.2.2: Choice Based Loss Severity Modeling Structure



CPB loans

The CPB portfolio is typically characterized by an affluent customer base (high net worth) who have had always demonstrated little to no sensitivity to economic cycles. As a result, the CPB portfolio does not have a discernible stress period. To provide some perspective, over the entire 2008-2009 historical stress period, the CPB portfolio witnessed only three loans defaulting and suffering losses. Taking into account the entire development period which spans from 2008-2017, there were only 300 CPB loans that experienced losses. Out of these 300 accounts that encountered losses approx~ 260 accounts pertained to the first liens and the rest ~50 accounts corresponded to the second liens. Given the limited delinquency history and corresponding losses during the 2008-2009 Financial Crisis and over the extended history, including both pre and post crisis timeframes which covered recent periods as well for the entire CPB population, there was a deliberate disposition to exclude these loans (both CPB first and second liens) from the LGD model development data pool, as there was no reasonable basis to build a model for such a sparse population of loans. Even if a model was built for the CPB loans, it would not yield statistically meaningful results resulting in lack of robustness of the modeling framework and unintuitive interpretation of the corresponding parametric specification.

However, being mindful of the feedback received from the model reviewers' on last year's model, a slightly different approach was utilized in the 2019 severity model to effectively model and forecast losses corresponding to the CPB portfolio. The strategy deployed to model the CPB loans slightly differed between the base and stress scenarios. For the first lien base scenario, an empirical lookup table was created leveraging the historical loan performance data from 201101 to 201706 to forecast the base losses.

Table 5.1.2.2 - Look up table of first lien base scenario loss

PRIN BAL	CLTV_MTM	Loss rate
1. <=300k		**0.53**
2. >300k	1: <=100	**0.32**
	2: 100-120	**0.33**
	3: 120~	**0.37**

In order to model the first lien stress scenario, a twofold approach was utilized. First, as a coping mechanism, (this helped remediate Limitation # 14920); the residential mortgage portfolio was chosen to serve as a proxy for CPB stress losses. The CMI residential mortgage portfolio was primarily chosen since it had a rich history of data with sufficient sensitivity to business cycles. The CMI residential mortgage empirical loss curve was plotted to gauge the level of peak losses. The empirical data revealed that there is usually a time lag in the loss recognition due to the delays in the foreclosure/bankruptcy processes. Therefore, while the historical stress period persisted from 2008-2009, the actual losses peaked from 2009-2011. This peak point of the loss curve was then utilized to calculate a Stress Multiplier. The multiplier was calculated by taking the ratio of the peak period losses to the recent period losses for the CMI residential mortgage portfolio. This multiplier was subsequently used to calculate the add-on losses over and above the base losses for the CPB first lien portfolio. Please note that the stress period is defined as 200901-201112 and the recent period is defined as 201507-201706.

Portfolio	N		PRIN BAL		Loss rate		Stress to Base Ratio
	Stress	Recent	Stress	Recent	Stres	Recen	

					s	t	
first lien Residential	40,713	1,932	7,337,575,063	211,619,458	0.50	0.41	**1.23**

For all model forecasting and backtesting needs, the derived LGD multiplier was utilized for all stress scenarios. For the baseline scenarios, the empirical lookup table was utilized to estimate the expected losses. Since there were only 50 accounts under second liens that experienced losses for the CPB portfolio, it was considered prudent to just assign full losses (loss rate=1) to these second liens for all forecasting and backtesting needs, including both stress and base scenarios.

Presented below are the specific loss treatments meted for specific loan categories.

Bankruptcy

While CAMU does not predict future bankruptcies, currently bankrupt loans receive the appropriate accounting loss treatment where collateral dependency is noted, as defined in the Mortgage Policy Manual (MPM) and Global Consumer Credit Fraud Risk Policy (GCCFRP).

Action Loan Treatment

CAMU refers to modified and sold loans collectively as action loans. CAMU can track modified loans past their modification date. However, CAMU has no way of tracking performance for sold loans once they are sold and their servicing is released. In the model described here, modified and sold loans are treated as censored observations in development. The CAMU Method A model suite deploys a 'Censoring' strategy to execute on the 'Action' loans. Through the years, CAMU had proposed several methodologies for the evaluation & associated treatment of these 'Action' loans. One important point to note at the onset is that 'Censoring' has no effect on the actual/post-action loan performances; it only affects the treatment of pre-action predicted probabilities, based on the type of approach used. Censoring methodology used in model development is to remove actual post-action performance for action loans.

The censoring approaches used by CAMU in the recent past ~~and as~~ described below ~~do~~ produce different predictions in backtest environment as they focus on addressing how 'action' loans should be treated in the backtest environment under each approach. Censoring does not impact the model forecasting results or the model's sensitivity to macro-economic factors as the model forecast always assumes that there are no future 'action' loans. This aligns with both the modeling objective and the future business plan which shows significant diminishing trends in both asset sales and loan modifications' in recent times. As part of CCAR annual planning exercise undertaken by Finance team at a Bank wide level, only the BHC Base scenario is allowed to assume prospective asset sales (i.e. only those deals not signed prior to the CCAR submission) – and this is done as an overlay/adjustment. NA Mortgage Finance confirmed that all other scenarios exclude any and all prospective sales at the request of regulators.

The 2017 Method A back test approach completely removed all 'action' loans from its purview resulting in significant limitations in terms of inconsistent population sizes between model development sample and backtest. The incomprehensive model performances are due to the removal of action loan pre-action performance and discrepancy in backtest volume across different snapshots compared to the general ledger.

To mitigate the aforementioned censoring limitation, for the 2019 CCAR submission, the Method A model adopted the censoring method for the treatment of the action loans. This method was deemed to be able to track and compare the actual and predicted portfolio performance up to the action date, which ensures backtest population completeness and consistency.

To be more specific, the censoring approach tried to mimic the model development data treatment and tracks the actual loan performance up to one month ahead of action date. Loan performance since the action date was not included. For example, if conducting a 27-month backtest for the Mar. 2016 snapshot, and a loan was modified on Dec. 2016, its actual performance history between snapshot and action date were included, but its delinquency status or loss on and after Dec. 2016 were removed.

As to the model prediction, it's important to note that the model is implemented in the Transition Framework, which forecasts the probability of a loan to be in defined transition and terminal buckets at each time. The Method A Residential Mortgage model suite is not designed to predict any future 'action' on loans and hence does not assign any probabilities for a loan to be 'actioned' on any given month. In this case, in order to be consistent with the portfolio actual performance, all the model predicted delinquency or termination probability and credit loss since the date of action were set to be zero. Presented below is a numerical illustration of the Naïve Censoring approach, executed this modeling process.

Naïve Censoring

As shown in the table below, the naive censoring approach evaluated the model without any adjustment on the "pre-action" raw forecast. The apparent advantage of this approach is that it is simple and straightforward. It doesn't require any adjustment to the model forecasted probability. However, it inadvertently leads to a misleading comparison between the actual and prediction. The rationale can be elaborated in a simplified example:

Suppose in a simplified situation where there are only two statuses of a loan, either terminated or still remain active, and suppose the model always forecasts 10% chance of loan termination at each month since snapshot, then at month t=1, model will predict $100\% - 10\% = 90\%$ of this loan to continue stay active, or say survival. In the next month, another 10% of the loan terminates, so the portion remain active becomes $90\% - 10\% = 80\%$. Going forward, the marginal probability of loan termination and active in each month can be written as below:

Raw Forecast	**t=1**	**t=2**	**t=3**	**t=4**	**t=5**	**t=6**	**…**
Marginal Termination	10%	10%	10%	10%	10%	10%	…
Active	90%	80%	70%	60%	50%	40%	…

Further suppose that this loan is actioned at month 5 (t=5), then the “Naïve Censoring” will remove both the termination and active probabilities since month 5:

Naive Censoring	**t=1**	**t=2**	**t=3**	**t=4**	**t=5**	**t=6**	**…**
Marginal Termination	10%	10%	10%	10%	.	.	…
Active	90%	80%	70%	60%	.	.	…

However, one can notice that we only removed the survival probabilities at the beginning of month 5 (60%). The other 40%, predicted pre-action liquidation probability, still remains in the backtest population. But it is known that this loan did not terminate before its date of action. By censoring the actual, we have removed 100% of this loan at month 5 whereas we only removed 60% of this loan when censoring prediction.

As shown in this simplified case, the Naïve Censoring removed **fewer** and always have **more** remaining in the population than in the “actual world”. And the difference is the predicted cumulative pre-action liquidation probability, for this particular loan, is 40%.

For additional details on the PCO logic, please refer to attachment 1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY.docx. Please note there is no empirical PCO distribution allocation over time as there are no actuals available for the PCO.

Walkthrough of Model Implementation and NCL (end-to-end) Loss Calculation

As part of CAMU’s commitment to address one of the MEA’s specific to this - 2019 Method A Residential Mortgage Model suite (Model Id # 167125) which required a walkthrough of the model implementation process with an overview of the end-to-end loss calculation, CAMU provided a single loan example demonstrating this Walk.

In order to illustrate the NCL calculation implementation, as an example, a deep delinquent loan was randomly selected from 201603 residential model backtest. Three months forecasted NCL and calculation related variables are pulled out from output file. A step-by-step verification process of NCL calculation logic implementation is presented in the attachment – ‘5.1.2 NCL_Walkthrough.xlsx’. All key logics comprise of each block of the calculation.

High Level Logic:

NCL = Severity Loss + PCO

Severity Loss = IVP*EAD*LGD

1. IVP - unconditional probability of charge Off

2. the component EAD * LDG is already calculated in severity model

PCO - Monthly Partial Charge Off Amount consists of four components -

1. Additional PCO – PCO happens loan in charge off procedure while BPO is reduced from previous month. The PCO calculation is based on 90% CLTV Policy

2. New bucket 7 – PCO happens when a loan newly entering bucket 7 or in charge off procedure with certain probability. The PCO calculation is based 90% CLTV policy

3. VP release - The PCO loss should be released when a loan is paid off. The release amount is based on probability of VP (slight adjustment, see the logic in Colum DD-DO) and Cumulative Partial Charge Off Amount up to this month

4. IVP release - The severity loss projects the life time loss for a IVP loan. The PCO loss should be released when a loan is completely charged off. The release amount is based on probability of IVP (slight adjustment, see the logic in Colum DD-DO) and Cumulative Partial Charge Off Amount up to this month

For all additional details, Please refer to attachment – '5.1.2 NCL_Walkthrough.xlsx'.

5.1.3 Model Segmentation

[Include rationale for portfolio segmentation and a discussion on how a particular methodology and model captures the key characteristics and the unique risk drivers of each portfolio segment.

If accounts are segmented into different groups for modeling, please provide evidence that the segmentation is of "appropriate granularity" with segments having sufficient data to develop robust and testable estimates that capture different underlying portfolio characteristics (like credit score ranges, or loan-to-value ratio ranges). Please also provide evidence that the segmentation logic is in line with the modeling objective. Demonstration that the segmented approach leads to better forecasting results (i.e. a smaller predictive error) must be provided in the model development documentation.]

The main objective of any statistical model is to produce the most accurate prediction for the dependent attribute. To achieve this, it is important to capture the model's heteroscedasticity across the granular segments.

The choice of desired level of model segmentation is critical in the development of consumer credit risk models. In some instances, the performance data is available only at the portfolio level, making this the most consistent choice for analysis. However, more often than not, historical performance data is available at the individual account level, including detailed information on the borrower characteristics, loan-level features and performance history. As a result, a lot of econometric modeling happens at a granular level to leverage the spectrum of information available.

Cohort level models produce more detailed delinquency and default risk estimates in comparison to aggregate portfolio level modeling due to its more detailed representation of heterogeneous portfolios. In a portfolio level model, the heterogeneity and specific idiosyncrasies of underlying granular segments composing this portfolio may be masked by the aggregate level statistics. The interaction of key risk factors may result in a higher (or lower) default risk and corresponding losses for a composite portfolio than that suggested by the granular segments when considered in isolation. Even for cohorts that are similar at the aggregate level, there are substantial and economically meaningful differences in credit quality between them. Existing research has attributed the differences between granular segment level modeling and aggregate portfolio level modeling to three main effects: simple cases of Jensen's inequality, aggregation bias, and masking of extreme observations. All of these findings have been analyzed, empirically tested and evidenced in the work of Chinchalkar and Stein (2010).

While the general consensus attunes to the fact that cohort level models are always better than aggregate models as they capture all granular level heteroscedastiBankes, there can be some disadvantages associated with granular level modeling:

1. If the granular modeling segments have an insufficient number of observations or lack an economic cycle to cover the complete macroeconomic lifecycle, these granular models might produce biased or undesirable results.

2. The constraints and underlying costs associated with granular level modeling in terms of implementation costs and/or space/resource shortcomings render the granular level modeling somewhat infeasible in terms of overall model execution.

At the end of the day, the optimal level of model granularity should strive to balance between model performance, data sufficiency and model execution in a timely, coherent manner.

Given these guiding principles and to add some perspective on what was achieved last year for segmentation; while the segmentation approach utilized for the 2017 Method A model suite were adequately justified based on the PD/LGD modeling framework, research on industrywide common segmentation approaches; limitations were still raised and recommendations were made to provide concrete factual evidence backing the 2019 segmentation scheme and also consider additional segmentation given the discernible differences in the underwriting/credit policies and borrower risk profile across business entities (CMI vs CPB vs US Retail Bank-Home Equity only) and other key portfolio risk drivers.

The approach utilized as part of the– 2019 Method A model suite encompassed the following ground rules while considering the appropriate segmentation scheme:

1. Selection of segments should align with commonly acknowledged business intuition. For example, there is a marked difference between CMI vs CPB specific originations (CPB entity typically cater to ultra-net worth clientele who often do not demonstrate any significant reaction to the upheavals in the macroeconomic environment or changes in the business cycles. In essence, any and all segments that justify a stand-alone analysis, backed by economic intuition, have been scrutinized within the scope of segmentation
2. Second, one of the key limitations cited on the 2018 CCAR models were the lack of supporting quantitative analyses backing the segmentation scheme. It is to be noted that while judicious care was taken in last year's modeling approach to implement the segmentation scheme based on all qualitative research on relevant cohorts such as channel, delinquency, etc., this year the

modeling team performed deep dive analyses on assimilating solid quantitative evidence that backed and corroborated on last year's model segmentation scheme.

3. Third, the objective of this year's segmentation scheme is not only to confirm and quantitatively attest to the segmentation strategy deployed in last year's modeling exercise but also identify additional new cohorts which have exhibited a propensity to demonstrate differentiated sensitivities across key risk drivers.
4. The model's finalized segmentation aligned with common business expectations and were further validated by statistical analyses that included decision tree and bivariate bubble plots.

5. The model's segmentation (encompassing the PD, LGD and DV Logic) are based on the entire development data that included both in-time and out-of-time samples.

Model	Development Period
Probability of Default Model	Feb-2006 – December 2017
Loss given Default Model	Jan -2008 - June 2017
DV Logic	Stress Period -Jan 2005 – December 2009[DRI data source] Recent(Non-Stress) – Jan 2010 - December 2017[DM data source]

Probability of Default Model

Keeping the given background in mind, the 2019 Method A Residential Mortgage PD model which utilizes a transition framework, improved upon the prior year's PD model segmentation scheme by incorporating supportive quantitative evidence with a focus on granular level details around the performance of key portfolio cohorts that correspond to business entities (CPB and CMI), origination channel (Retail vs Broker vs Third Party), loan delinquency levels and other key risk indicators.

The segmentation framework, entailed a four (4) step stepwise execution strategy, which is outlined as follows-

1. **Decision tree analysis**

The modeling team utilized the decision tree engine for the 2019 modeling process which helped identify additional key modeling segments or conform the pre-identified segments from the prior model suite. The tree analysis leveraged statistical modeling to generate an inverted tree which ranked the segments by order of their statistical significance and relative importance to the target metric. For the residential mortgage portfolio, decision tree analyses were executed on some of the important transition equations. The importance of the transition equations are determined by the size of the population cohorts they were being modeled on and their relative effect on overall model performance. The target metrics were defined as the roll-forward delinquency or terminal (IVP, VP) loan states.

Advantages of Decision Tree:

Leveraging the decision tree engine into the realm of credit risk profiling has many advantages –

- Decision trees are a form of multiple variable (or multiple effect) analyses. All forms of multiple variable analyses allow us to predict, explain, describe, or classify an outcome. This multiple variable analysis capability of decision trees enables to go beyond simple one-cause, one-effect relationships and to discover and describe things in the context of multiple influences.
- Provides an objective overview of analytically validated data to back up commonly known business decisions/ economic intuition.
- The ability to analyze thousands of data points and create appropriate granular level cohorts/nodes: there is no limit to the amount of columns or rows of data/business information that can be leveraged.
- Easily slice and dice customer data into infinite number of sub-groups/cohorts based on the PD/LGD model design and the underlying credit policy/risk management processes: the segmentation allows to easily re-run data and categorize them into sub-segments (either user-defined or auto-generated) to focus on particular risk - attributes of a given group.
- Correlate and predict behavior across sub-groups/cohorts more accurately (information) that can help uncover relationships and shared key risk drivers attributes across specific segments which can improve the model's fit through the introduction of important interaction variables.

For the Method A model suite, CAMU conducted a decision tree analysis to determine the appropriateness of segmenting the risk portfolio into Modified vs Non-Modified loans and by product types (Residential vs HELOC vs FRHEL). In this analysis the entire model development panel dataset pulled on the risk portfolio (Modified, Non-Modified RM, Non-Modified HELOC and Non-Modified FRHEL) was used in the creation of the tree. The target variable/metric was "IVP/VP status within 27 months of file_dt". The input variables considered in the segmentation analysis included:

- Portfolio indicators such as MOD, non-mod RM, non-mod HELOC and non-mod FRHEL
- Delinquency status of the observation month
- Delinquency history variables such as number of 30/60/90 days past due in the past 6 months
- Business entity such as CPB or CMI
- Origination channel such as Retail, Broker, Correspondent and Wall Street
- Lien position
- Product type such as Fixed rate (FRM), Adjusted rate (ARM) or interest only (IO)
- Loan age

- Loan type such as FHA, VA or Conventional
- Origination type
- Property type
- Loan purpose
- Documentation Type
- Loan Purpose

All of these proposed attributes have been carefully selected based on prior year's model results, business intuition and model reviewers' feedback. An important point to note here is the fact that there was a deliberate attempt to not introduce any continuous variables as part of the decision tree analysis since addition of continuous attributes often leads the decision tree to loose information when it tries to categorize (step-up/step-down) these variables in different categories.

Please see tab 'IVP tree' and tab 'VP tree' in attachment '5.1.3 DecisionTree_RM_PD.xlsx' for the illustration of all the relevant tree analyses.

1. Delinquency status

2. Lien Position

3. Origination Channel

4. Age of Loan

5. Portfolio Segment (MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL)

It is not entirely surprising to witness loan level delinquency status as the most important predictor. This conforms to the transition framework of the model, which naturally segregates the loans by their delinquency status. While lien position and portfolio segment (MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL) have already been in use in prior segmentation scheme, the decision tree analysis conducted this year, established the importance of origination channel as one of the key risk indicators that influence the loan performance.

In addition, as shown in the IVP tree, for those 90+DPDs delinquencies (MBA_DELQ_STATUS_CD in 4, 5, 6,7), the most important split is to segment the population by MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL.

Similarly, within the VP decision tree, the following list illustrates the chronological order of the attributes within the tree-

1. Ever 30days past due

2. Portfolio Segment (MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL)

To summarize, the importance of portfolio type separation in both the IVP and VP decision tree analyses provided concrete statistical evidence, which validated the model's existing scheme of segmenting the portfolio into MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL and developing separate models for each of them. Further the continued appearance of the delinquency status, justifies the rationale for continuing to segment the model by the delinquency buckets. All other nodes, which came up within the top five ranking were effectively, used in the model as either intercept or interaction variables. To this end, for the 2019 Method A model suite, the finalized model segmentation scheme segments the overall portfolio by delinquency status and portfolio types (RM MOD vs RM Non MOD). A separate model has been built to model the modified loans of which the residential mortgage modified loans is modeled as separate segment. More details on the MOD portfolio segmentation can be found in Section 5.1.3 of the MOD MDD.

Additionally, CAMU also conducted segmentation analysis based on the panel data corresponding to the non-modified residential mortgage loans only. Since the Method A PD model is estimated as a set of equations which predicts the transition probability from one-status in a given observation month to another status in the following month; as part of supplemental segmentation analysis, CAMU chose the target variable as "whether the loan rolls into a deeper delinquent buckets in the next month". Please see tab 'RM_Worse_ tree' in attachment '5.1.3 DecisionTree_RM_PD.xlsx' for the illustration of all the relevant tree analyses. In this "roll-to-worse" tree, the following segments were deemed important based on the tree-analysis.

1. Delinquency status as of the observation month
2. Historical Delinquency information
3. Loan Age
4. Loan Type (Conventional vs Government)

The importance of the delinquency status at the observation month indicated and confirmed that it's appropriate to develop a set of one-month-ahead transition equations by different starting buckets as has been always done in the previous version of the Method A PD model.

It is important to note here that although CAMU did include a business entity indicator such as CPB loan indicator into the decision tress analysis, it did not appear as an important node in the decision tree analysis. Given this factual quantitative evidence, CAMU considered it prudent to not develop a model specific for the CPB portfolio.

Hence, instead for creating separate equations solely for the CPB portfolio, the CPB loans have been combined with the CMI portfolio in the model estimation process as the appropriate coping strategy. Additionally, a dummy indicator for the CPB loan was tested as part of the model development process to ascertain that the model was able to discern and isolate the transition rates for CPB vs CMI loans, keeping all other explanatory variables within the modeled equation constant.

Further, during the Method A PD model development process, CAMU performed deep-dive analysis that compared and contrasted the model's performance across macro-economic attributes, key loan level risk factors and business cycles. This tool was extensively leveraged by CAMU to examine the model's one-month-ahead predicted roll rate versus the actual model fit for important portfolio all modeled segments, such as CPB and CMI, Conventional and Government loan, FRM and ARM. (Please refer to the variable selection process in Section 5.1.4 for pertinent details).

Based on the analysis performed, if differentiated sensitivities were noted for different portfolio segments (such as - CPB vs CMI loans) to changes in the macroeconomic attributes, interaction variables between macro-economic attributes and these segments were introduced and/or other key loan level driver specific interaction terms were tested in model estimation and correspondingly applied to the modeled equation, as deemed necessary (please refer to the variable selection process in Section 5.1.4 for pertinent details). The introduction of such interaction variables had two fold advantages –

1) Captured CPB specific sensitivity to key risk drivers;

2) Maintained the model's simplistic structure without the need to add any more equations, while still preserving the model's performance and overall robustness performance

To summarize, the importance of portfolio type separation in both the IVP and VP decision tree analyses provided concrete statistical evidence, which validated the model's existing scheme of segmenting the portfolio into MOD/Non-MOD RM/ Non-MOD HELOC/Non-MOD FRHEL and developing separate models for each of them. Further the continued appearance of the delinquency status, justifies the rationale for continuing to segment the model by the delinquency buckets. All other nodes, which came up within the top five ranking were effectively, used in the model as either intercept or interaction variables. To this end, for the 2019 Method A model suite, the finalized model segmentation scheme segments the overall portfolio by delinquency status and portfolio types (RM MOD vs RM Non MOD). A separate model has been built to model the modified loans of which the residential mortgage modified loans is modeled as separate segment. More details on the MOD portfolio segmentation can be found in Section 5.1.3 of the MOD MDD.

2. Bivariate-analysis

With the execution of the decision tree analysis, once the key segments were identified, an exhaustive bivariate analysis was subsequently conducted on these key segments to understand the behavior of these key segments across business cycles or key macro-economic variables. The bivariate analysis under the Method A modeling exercise helped in two distinct ways--

a) First, the bivariate analysis aided with understanding the different relationship between loan behavior and key attributes such as HPI, mark-to-market CLTV for different portfolio segments. The bubble plot created for bivariate-analysis also helped to determine if there is any nonlinear relationship between the delinquency, default or prepayment risk and those key attributes. If deemed necessary, appropriate splines will be created for each segment and tested in model specification.

b) Second, it justified the use of interaction variables, which improved the model's fit without increasing the model's complexity. For instance, there are some loans which were originated through the Wall Street channel pre-crisis (please note there have been no new Wall Street originations post crisis). Common intuition dictates that these loans are inherently riskier since they were originated pre-crisis through sub-optimal underwriting standards and via a third party originator. However, a bivariate analysis shows that for the Wall Street Channel, the delinquency transition risk decreases steadily when month-on-book (N_M_MOB) increases. This indicates that the distinction between the Wall Street and non-Wall Street diminishes as a loan ages. Further, given the successful track record of these existing Wall Street loans in terms of their performance and their resilience to crisis, it was considered prudent to not over-penalize these loans and instead use a time-decay factor as interaction to decay their performance over time.

For additional details on the bivariate analyses conducted this CCAR cycle, please refer to Section 5.1.4 of the MDD.

3. Regression Analysis

Once the bi-variate analysis was conducted and performance of key segments augmented through the introduction of splines and/or interaction attributes, regression analysis was conducted whereby the model's both in-sample/out of time sample performances was examined not only at the total portfolio but at the granular segment levels too. The regression model results were used to further fine-tune the model specifications and intuitively measuring the add-on effect of specific attributes. Further the end-to-end backtesting framework also had built – in granular segments within its reporting structure for evaluating the model's backtest performance at key segment levels. For additional details on the regression analyses results, please refer to Section 5.1.4 of the MDD.

4. Characteristic Analysis

The finalized segmentation scheme, as proposed above has been further validated based on end-to-end model results which demonstrate correct risk-ranking and accurate segment level results as part of model's overall fit and performance.

DV Logic Segmentation

The rationale justifying the segmentation of loans into prior distressed valuation (PDV) without prior distressed valuation (NPDV) is based on data availability of distressed properties and the chosen modeling methodology. For all properties that had a prior distressed valuation, it is considered best practice to leverage this value for accurate assessment of the property. However since not all properties have a prior distressed value available, the workaround is to impute a value based on the original home value. This justifies the first segmentation into NPDV and PDV segments.

Second, based on a comprehensive assessment of all relevant loan level attributes that affect a property's distressed value across stress and non-stress periods, owner-occupancy and lien position came out as significant parameters for the NPDV segment.

The attachment –'5.1.2 pivot_NPDV_Haircut_newData_HPIV4byQtr.xlsm' shows the haircut ratio across time for the NPDV segment. It shows among all the dimensions that were tested, only occupancy type and lien position came out significant. Haircut ratio is significantly lower on the owner-occupied property than the non-owner occupied property, and it is also lower on the second lien when compared to first lien. Hence, it was considered prudent to segment the DV haircut ratio based on these two attributes for the NPDV segment.

The attachment –'5.1.2 pivot_PDV_Haircut_newData_HPIV4byQtr.xlsm' shows the haircut ratio across time for the PDV segment. It shows no significant difference in all of the dimensions that were tested. Hence, it was considered prudent to not create any further segments for the PDV segment.

For additional details on the DV Logic segmentation differences, please refer to Section 5.1.2 and Sections 3.1 & 3.2 of the MDD.

Loss Given Default Model

The 2019 Method A LGD model suite also utilized a similar approach towards segmentation, as the PD model. To provide some context, the Method A LGD model redeveloped its modeling methodology to model loss severities based on loss types (partial, zero, full) compared to the prior approach which modeled the losses by their disposition types (3F, CO, SS, RO). The rationale behind the change in the modeling approach was guided on a needs based approach to switch away from a methodology that tried to model loss outcomes which were predetermined by strategic business initiatives, instead of the regular BAU process. Similarly, the old exclusion criteria that excluded all zero loss loans from the development data set, was also considered inappropriate, given the influx of zero loss loans in recent time periods due to improving housing prices. All of these rationales have been exhaustively discussed in Section 3.1 – Modeling Approaches of this MDD.

Given this context, the first natural segmentation for the LGD model pertained to segmenting the losses by their occurrence type (full vs partial vs zero). This segmentation was further validated by a separate bi variate analysis, which examined the bivariate relationship between the loss severity and the mark-to-market LTV, aka borrower's LTV which is considered a key risk factor of the LGD model. Intuitively speaking, higher LTV is associated with higher severity rate due to relatively higher debt compared to the total property value. To provide some background around this analysis, the prior modeling approach experienced a non-monotonic relationship between the MTM LTV and loss severity at lower LTV bands, i.e. loans with LTV < 100 experienced higher loss rates compared to loans with LTV = 120, especially for the first lien CO and SS disposition types. With no option to alleviate this V-shaped anomaly, the modeling team deliberately changed the modeling approach and implementation the aforementioned segmentation to correct the rank ordering (monotonicity) of the loss severities to the current LTV. For additional details, please refer to Section 3.1 of the MDD.

At the next step, a decision tree analysis was conducted to identify the key segments for the LGD model. Please refer to attachment – '5.1.3 Severity_DecisionTree.xlsx' for additional details. Based on the decision tree analysis, the following nodes came up in the chronological order of importance-

a. Lien Position – Portfolio [First vs Second vs VA]
b. Property State
c. Business Entity (CMI vs CPB vs US Retail Bank)
d. Loan Purpose (Purchase vs Cashout vs Refinance)
e. Origination Channel
f. Occupancy Type

Given the importance of 'Portfolio/Lien Position' and 'Business Entity' ranking among the topmost nodes, separate models/treatments were developed for first, second, VA and CPB loans. This segmentation conformed to business intuition too as VA loans have unique loss identification policies dictated by the guaranty clauses that are associated with the VA loans. The remaining attributes were added as predictors within the modeled equations. The finalized segmentation scheme pertaining to the 2019 Severity model segmented the overall loss portfolio by their loss type and their lien position. Different treatments were employed for the VA and CPB loans to estimate their losses separately.

Similar to the PD model Decision Tree and segmentation analysis, bi-variate, regression and characteristics analyses were all conducted based on the relevant segments of interest for the Severity model too. Results from the bi-vicariate and regression analysis can be found in Section 5.1.4 of the MDD. Characteristics analyses results are present in Section 6.3 of the MDD.

To summarize, the 2019 Method A model development process executed and delivered a more refined segmentation analysis that leveraged quantitative statistical tools to discern the segments of interest based on their sensitivities across key risk drivers. These segments were further tested across business cycles and other broader economic factors to fine-tune the model's performance for these specific segments. Results from the segmentation were eventually utilized in the regression analysis, which itself underwent an exhaustive iterative analysis before the modeled equations were finalized. For the attributes that didn't turn up as the top nodes within the decision tree engine (such as FICO, LTV, ARM vs FRM) to merit separate modeled segments, these were incorporated as interaction variables to capture their differentiated sensitivities. Further the model's characteristic's analysis provides insights on the model's performance along these risk drivers.

- Was segmentation explored and if no segmentation is possible, appropriate rationale/documentation should be provided?

  Yes, segmentation was explored as part of the modeling process. The finalized segmentation scheme was based off a statistical decision tree and accompanying deep-dive analysis at granular segments.

- What tests were performed to confirm the segmentation scheme meets the modeling objective and/or generates better forecasting results?

  Quantitative statistical analysis (decision tree) was conducted this modeling cycle to conform the existing segmentation scheme and identify any new segments of interest for both the PD and LGD models. Please see response above for additional details on the decision tree analysis.

- What were the results of the segmentation tests?

  The segmentation testing helped validate the prior model's segmentation scheme with quantitative, factual evidence and aided in exploring any new segments of interest. The addition of the quantitative analysis helped address last year's model limitation that was centered around the prior segmentation being grounded on only qualitative judgment. Please see response above for additional details

- How were the test results interpreted to refine the segmentation?

  The results of the decision tree analysis were closely reviewed by the model developer in order to discern the segments of interest based on their sensitivities across key risk drivers. These segments were further interpreted across business cycles and other broader economic factors to fine-tune the model's performance for these specific segments. Results from the segmentation were eventually utilized in the regression analysis, which itself underwent an exhaustive iterative analysis before the modeled equations were finalized.

- What was the final segmentation scheme used?

  The Residential Mortgage PD model segmented the portfolio by delinquency status and then product type (residential mortgage vs FHREL vs HELOC). The LGD model segmented by lien position. All appropriate rationale underlying the chosen segmentation scheme has been discussed above.

  - If segmentation is based on business judgment, appropriate rationale must be provided to support the segmentation scheme.

    The segmentation scheme used in residential mortgage model suite is based off a comprehensive rationales of business purposes, economically intuitive risk drivers, and statistical testing results which included the decision tree, bivariate sensitivity analysis, regression analyses and lastly the characteristic analyses that validated the model's end-to-end results for the chosen segmentation scheme.

- What is the evidence of segmentation optimality? How do the optimization criteria relate to the overall aim of the model?

  The model results are reviewed and analyzed on overall level and granular segment levels level. The overall aim of the model is to forecast portfolio level gross and net credit losses, delinquencies, default and voluntary prepay for residential mortgage, loans which is consistent with our overall proposed segmentation scheme.

- Is the segmentation technique applied robust over time (i.e. is it invariant to the starting point)?

  Yes, the segmentation technique is robust over time, invariant to the starting observation point.

- Are the segments internally homogeneous, and do they differ significantly between each other? What measure is used to determine homogeneity / heterogeneity of segments?

  The segments differ significantly between each other due to the nature of different mortgage lending products (RM vs HELOC vs FRHEL) and corresponding credit loss policies (Government Insurance vs Conventional), which are described in details above. Within each segment, loans share similar sensitivities to key risk drivers.

- Was segmentation addressed through segment-based interaction variables instead of creating separate model segments? If so, provide statistical evidence that variables not used in the interaction terms show similar effects across segments or otherwise eliminate the need to build models for separate segments.

  The decision tree analysis and discussions presented in preceding sections describe the approach taken to choose between segmentation vs interaction variables.

5.1.4 Model Variable Selection

[Documentation should include the entire variable selection process; details on variable selection technique used along with the justification for explanatory variables selected, including coefficients from statistical models, measures of their statistical significance, and qualitative assessments where appropriate. Where relevant, descriptive statistics, including mean, median, minimum, maximum, and standard deviation should be outlined.

[Documentation should include the entire variable selection process; details on variable selection technique used along with the justification for explanatory variables selected, including coefficients from statistical models, measures of their statistical significance, and qualitative assessments where appropriate. Where relevant, descriptive statistics, including mean, median, minimum, maximum, and standard deviation should be outlined.

Apply consistent variable selection approaches across model segments or modeling system component models. Pay careful attention to the consistency of testing and approaches for selecting variables and how macroeconomic effects are incorporated in models.

Identify any use of variables not forecasted by regulators. The inclusion of such variables in models carries additional analysis considerations, including back-testing exercises and on-going monitoring to assess the model risk associated with forecasting these variables for model use. Variable testing and selection documentation should be well-organized and clearly presented so that it can be leveraged by others to develop a robust assessment of variable consistency across suites of models and effects and testing approaches across businesses, regions, and products.

After the modeling data has been collected and validated, the modeler should engage with the business experts to develop a list of potential drivers that are meaningful from a business perspective and can logically explain movements in the dependent variable; the directional impact of each driver; and the potential importance of each driver relative to the others.]

Variable selection for any econometric model involves selecting a set of covariates that can help construct the 'best' model. A decision to keep a variable in the model might be based on common business intuition/qualitative evidence or statistical significance or both. The criteria for inclusion of a variable in the model vary between the overarching objective of the model and the data/information available to build the model. The common approach to statistical model building is minimization of variables until the most parsimonious model that describes the data is found which also results in numerical stability and generalizability of the results. Some methodologists suggest inclusion of all relevant variables of interest in the model regardless of their significance in order to control for confounding effects of the model. This approach however, can lead to numerically unstable estimates and large standard errors. In addition, it may also lead to a model with unexpected effect of the independent variables. Hence, it is very important to conduct a thorough and purposeful selection of variables in regression methods to construct a model, which is robust over time. Further, the chosen approach should also be consistent across model segments/equations and across businesses and products.

As the first step towards the variable selection process for the 2019 redeveloped Method A Residential Mortgage model suite, CAMU reviewed all the feedback and limitations (both technical + functional) posed on the prior model suite to assess the scope of its prioritization efforts for the 2019 model redevelopment process, dependent on resource/technological constraints and the submission timeline. The 2018 model validation efforts also included a review of the Method A production code by the model's functional reviewers. Given the IRMO Code review results, CAMU extended its redevelopment efforts to include enhancements to the model code, in particular to the PCO logic, amortization logic for residential mortgage products and the way the CLTV spline attribute is defined to make it consistent between model development and model implementation. Additionally, as part of self-identified model/process improvement initiatives, CAMU identified and proposed to unify the product definition across businesses and all pertinent model usages to aid in more accurate and consistent interpretation of model results.

Once the key redevelopment areas were identified for the 2019 CCAR model development process, CAMU presented the 2019/2020 NA Mortgage CCAR model redevelopment plan with the Model Risk Management(MRM) team to make sure all the key model stakeholders were well aware and approved the 2019 modeling objectives. For more information on the 2019 CCAR Modeling Plan and the associated meeting minutes, please refer to attachment '5.1.2 2019 CCAR Modeling Plan_20180619.pptx'.

The following bullets represent the itemized areas of redevelopment for the Method A residential mortgage model suite for the 2019 CCAR process.

1. Quantitative evidence supporting the segmentation scheme

2. Separate evaluation of the modeled vs non-modeled equations using a waterfall logic that is based on limited and volatile data trends and statistical testing. All transitions go through a waterfall logic wherein they are tested based on their total overall volume, significance of C-statistics and their contribution volume to the source/destination transition cells and underlying volatility within these transitions. All transitions, which passed the waterfall, were modeled while a lookup table created for stress and non-stress, respectively across some critical segments (In-Trial vs not –in-Trial, Govt vs Conventional) for the non-modeled transitions, based on the model's in-sample data. Further the non-modeled transitions were shocked using a 10% rate [based on the Model Testing Guidance] to validate the stability of assumptions to shocks

3. Improved model's sensitivity to recent and stress periods through the introduction of new macro-economic variables to capture equity market performance and is cascading effects on real estate performance

4. Introduced new HPI burn-in and VP interest rate burnout effect to capture portfolio composition changes in response to the changes in the broader macro-economic environment

5. Deep-dive analysis was conducted and variable transformations (change of change, interaction) were introduced to capture the model's stress peak losses and improve the model's VP equation sensitivity to macro-economic factors and to capture the model's recent period BUK7 -> IVP increase.

6. Enhanced amortization balance logic and partial charge-off calculation for estimating losses

7. Included recent period data for both the PD and the loss severity models. The inclusion of the recent data points helped with capturing the recent portfolio trends and origination profile(s).

8. The LGD modeling framework was changed from modeling losses by their disposition types to modeling losses by their outcomes (Full vs partial vs zero) for first lien loans. This change in modeling framework had a three-fold advantage – 1) Zero loses were no longer excluded from the model development sample, especially given the influx in zero losses in recent times due to consistently increasing HPI,2) Removed the uncertainty associated with modeling losses by their disposition types, which are strategic decisions that cannot be predicted in advance, and 3) aligned the modeling approach across first and second liens for a more consistent interpretation and comparison of model results.

9. Lookup LGD logic for CPB loans and application of VA guaranty amounts to the calculation of VA losses

10. Leveraged reconciled CPB loss data for model redevelopment

Given the targeted model redevelopment goals, the 2019 Method A model variable selection process started fresh for the RM PD model with no reference to prior model's variable selection processes.

PD Model Variable Selection

The main objective behind any variable selection process is to select the 'best' subset of predictors that explains the model well. Variable selection constitutes a critical aspect of any econometric modeling process for many reasons, including the important ones that are listed below-

- The main purpose of any modeling exercise is to explain the data in the simplest way, without much redundancy. According to Occam's razor theory, when presented with several plausible explanations, one should select the answer with the fewest specifications. In the context of credit risk modeling, this implies that the parsimonious model that fits the data well is the best.
- Adding too many predictors can add noise to the estimation process, impacting the credibility of the estimation. The model's degrees of freedom are also wasted
- Multicollinearity is caused by having too many variables trying to do the same job.
- Removing the redundant parameters can reduce the costs as well as the efforts spent on running the model
- Furthermore, it is very important that that variables selected into the final model are having expected signs so that the relationship between the dependent and independent variables can be well explained by business and economic intuition.

The next set of steps lists out the chronological order of the variable selection that was attempted on the Method A model suite.

Variable Selection (VS) Steps 1 and 2

VS 1. The first step included collecting all important candidate variables (loan level static and dynamic attributes), bureau attributes and macro-economic parameters as supplied by the GCRM team.

VS 2. The next step involved understanding and re-evaluating the nuances around the transition framework itself. Under a transition framework, roll –rates can be classified as either rare or non-rare. Rare transitions usually identify the jumping transitions where a loan skips transitioning to or back to the next adjacent cell and essentially jumps through a few cells. These jumping transitions were termed as non-modeled transitions in the prior model suite. For the 2017 PD model, there were thirty-five (35) non-modeled transitions. These included the following roll rates-

1) Jumping transitions from current (or low delinquent) to deep delinquent or IVP
2) Transitions from deep delinquent to VP
3) Partially cured transitions from deep delinquent to low delinquent without full cure

For each of the afore-mentioned non-modeled transitions, as part of the 2017 CCAR modeling process, CAMU derived a historical average roll rate without separating the stress period and the recent period and applied this constant number for the entire historical time horizon as well as future forecasting.

While using a constant number for non-modeled transitions aided in simplifying the model's complexity, the model's fit can be potentially sacrificed across time periods, specifically across stress and the recent periods with this approach since macro-economic conditions and the actual roll rate differ vastly across these periods. Given this limitation, the modeling team conducted a deep-dive analysis this modeling cycle to understand the nuances and the associated volume around some of the rare – transitions.

The following observations were noted as part of this analysis-

- As already mentioned in prior sections of the MDD, the rare transitions, in particular the jump to worse ones are typically low volume transitions in recent observations. Bank's portfolio today is characterized by low default and loss rates. A slightly biased transition rate can lead to significantly different model results for such a portfolio. Hence it becomes extremely important for these transitions to be as accurate as possible especially for most recent time periods.
- Second, these transitions are also extremely volatile over time with some of them exhibiting significant differences between stress and recent periods. Therefore, using a singular averaged constant roll rate over the entire time horizon cannot fully differentiate the extreme differences in actual performances over time.
- Furthermore, some transitions did not align with the macro-economic trend/movement, and performance in the historical stress period was better than in recent periods, thus rendering it almost impossible to build a statistically meaningful and sound model.

Given the above challenges as a precursor to this year's model development work and the cited limitation on the model's complex framework, CAMU laid out a stepwise logic/criteria to completely reevaluate the modeled vs non-modeled transitions and tested out several alternatives(cited in Section 3.2.2 of the MDD) to this regard, before finalization of the chosen approach.

The 2019 CCAR model development process introduced a standardized process to determine modeled vs. non-modeled transitions and the non-modeled transitions value settings. As iterated before, this process was introduced explicitly to address the model limitation cited by model reviewer last year regarding: 1) lack of a standard process for determining non-modeled transitions and their values; 2) as a result, the overly simplified non-modeled transition assumption is one of the main factors to model's recent period over-estimation.

In last year's model development process, the selection of non-modeled transitions was mostly based on volume and model developers' judgment with the value simply set to the long-term historical average. The shortcomings of this process are rather obvious. First, some selected non-modeled transitions are actually intuitive transitions with sufficient volume or impact, and potentially model-able with intuitive explanatory variables, such as the curing transitions of BUK5->BUK2, BUK5->BUK3 or BUK6->BUK4 etc. Second, the long-term historical average could be significantly biased towards historical portfolio performance due to Bank's continuous efforts to improve portfolio quality and organic growth.

To address these limitations and further improve the modeling process, CAMU has developed a standardized approach to select model-able and non-model-able transition cells and applied consistently across models. The approach is based on economic intuition, volume, and contribution to the source cell and destination cell as described in preceding paragraphs of the MDD. An attachment [5.1.4 RM - Modeled vs Non-Modeled Transitions.xlsx.] detailing the determination process of the modeled vs. non-modeled transitions has been referred and described in the following sections and also attached for model reviewers to understand and replicate the selection process. CAMU is prudent in selecting transitions that cannot be modeled. As a result, a few curing transitions as mentioned before become model-able, such as BUK5->BUK2, BUK5->BUK3 or BUK6->BUK4 etc. following the waterfall approach in order to improve the model's accuracy and sensitivity.

Transitions were evaluated based on following waterfall criteria, which encompassed the business rationale, transition volume, volatility and contribution amount (source or destination cells) and model soundness (C-statistics). To determine modeled versus non-modeled transitions, the entire dataset from 02/2006 to 12/2017 was leveraged which included the development sample, in-time hold out validation sample as well as the out-of-time validation sample. The logic is outlined as below-

1. **Not Aligned with business expectations**. The first rule of the waterfall logic considers commonly observed business practices to filter out intuitively impossible transitions such as the jumping-to-worse transitions (e.g., BUK1-BUK6 or BUK7).

2. **Not correlated with macro-econ trend** in an economically, intuitive way. Typically, historical evidence suggests that stress roll to worse rates should be higher than the non-stress roll to worse rates.
3. **Limited volume:** on average less than 10 loans/month transitioning from the source delinquency bucket to the destination DLQ bucket
4. **Limited contribution portion:** The number of loans transitioning represent less than one percent (1%) of the total volume for either the source or destination bucket. Any further granularity will lead to an increase in volatility and overfitting of the applicable data.
5. **C-statistics < 0.6** (the variable selection criteria as per MRM's model testing guidance). The model type is logistic regression and all pertinent model diagnostics results are shared in Section 6.1.8.

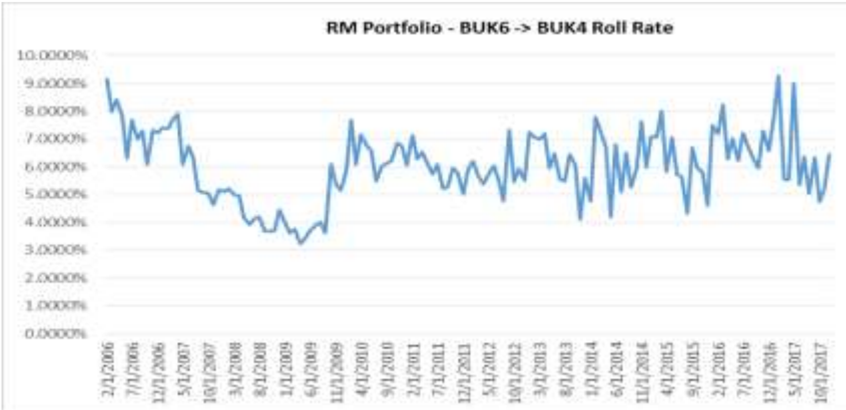
Given this waterfall structure, total of twenty-eight (28) transitions were deemed eligible for being modeled as separate equations and the remaining twenty-eight (28) equations were deemed as non-modeled transitions. For the detailed modeled vs non-modeled transition design in 2019 Method A RM model, please refer to Figure 5.1.4.1 in which the pink cells are non-modeled transitions; the white cells are considered the reference cells and the green cells represent the modeled transitions'.

The 2017 Method A first lien PD model had eight transitions which were deemed as non-modeled transitions as part of the 2017 CCAR process. Based on the new 2019 Waterfall logic these eight transitions were deemed eligible to be modeled as separate equations in order to improve the model's accuracy and sensitivity.

These eight transitions corresponded to BUK5->BUK2, BUK5->BUK3, BUK6->BUK4, BUK7->BUK01, BUK4->IVP, BUK5->IVP, BUK3->VP and BUK7->VP equations respectively. Please note that all of these eight transitions which were non-modeled in the previous model and are classified as modeled equations this time due to the use of the waterfall approach, DO satisfy either business justification and/or align with macroeconomic trends, as iterated below.

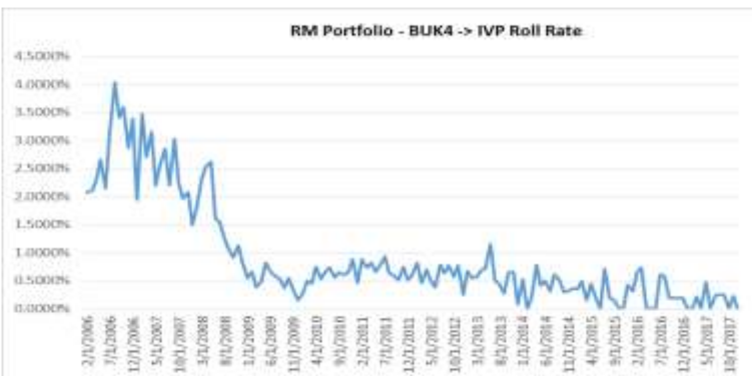
BUK5 - >BUK2/BUK5->BUK3/ BUK6->BUK4/ BUK7->BUK01

Business Justification – This is not an uncommon transition as there might be borrowers who might have received a sudden influx of cash(year-end bonus, lottery, monetary gifts, etc.) who use the extra cash to make payment and partially cure/fully cure. Also note that transitions also align with macro-economic intuition, as an example, the BUK6-> BUK4 roll rates are lower in stress compared to recent times, indicating lower cure rates during stress.



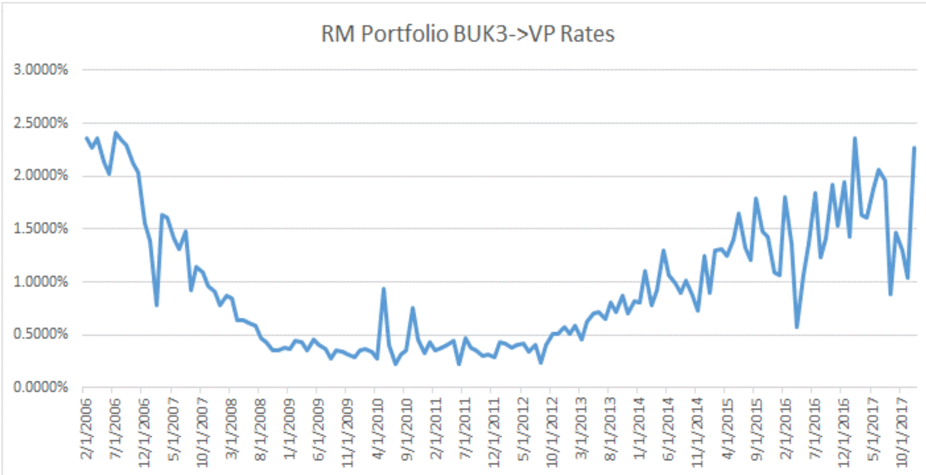
BUK4 - >IVP/BUK5->IVP

Justification – It is not uncommon to come across certain loans which are charged off at BUK4/BUK5 based on the prevailing business policy and the concerned borrower's presumed riskiness. Also these transitions align with macroeconomic trend(using BUK4->IVP) as an example, roll to IVP are much lower in recent times due to stringent underwriting criteria and stronger credit profile of the borrower.



BUK3 - >VP/BUK7->VP

Business Justification – These are again not uncommon transitions as there might be borrowers who can sell their homes , thus fully paying off their loan and move to a different city for employment/other specific relocation needs. As shown in the Figure below(BUK3->VP used to illustrate the point), the actual roll rate drops during stress period which experiences stressed economic conditions in terms of decreasing HPI, higher unemployment, lower GDP/income, etc. and improves during non-stress/recent times that reflects stronger economic growth.



Additional details, can be found in attachment mentioned in Section 6.1.9.

Please see Figure 5.1.4.1 and Table 5.1.4.1 respectively for the modeled vs non-modeled transition segregation and the logic that underlines this segregation. Overall, this structure has two-fold advantages-

a) First and foremost, it created a logical approach to logically and statistically separate the modeled vs non-modeled transitions
b) Second, it helped to simplify the model's complexity and assisted in justifying the logic behind each and every modeled equation

Figure 5.1.4.1 RM Transition Matrix Modeled vs Non-Modeled Equations

	BUK1	BUK2	BUK3	BUK4	BUK5	BUK6	BUK7	IVP	VP
BUK1	Ref	Model	Non-Model	Non-Model	Non-Model	Non-Model	Non-Model	Non-Model	Model
BUK2	Model	Ref	Model	Non-Model	Non-Model	Non-Model	Non-Model	Non-Model	Model
BUK3	Model	Model	Model	Ref	Non-Model	Non-Model	Non-Model	Non-Model	Model
BUK4	Model	Model	Model	Model	Ref	Non-Model	Non-Model	Model	Non-Model
BUK5	Model	Model	Model	Model	Model	Ref	Non-Model	Model	Non-Model
BUK6	Model	Non-Model	Non-Model	Model	Model	Model	Ref	Model	Non-Model
BUK7	Model	Non-Model	Non-Model	Non-Model	Non-Model	Non-Model	Ref	Model	Model

Table 5.1.4.2 below presents the number of non-modeled transitions that were determined by the waterfall.

Table 5.1.4.2 Non-Modeled Transition Waterfall Summary

	# of Non-Model Determined
Logic 1 & 2	17
Logic 3	2
Logic 4	9
Logic 5	0
Total	***28***

a) For the details on how non-modeled transitions were determined, please refer to the attachment '5.1.4 Residential Mortgage - Modeled vs Non-Modeled Transitions.xlsx.
b) Given the first logic which relates to non-alignment of the transitions with the prevailing business expectations/risk driver and/or unintuitive macro-economic trend, a total of seventeen (17) equations were deemed as non-modeled given these two logics of the waterfall. All remaining transitions flowed to Logic # 3 of the Waterfall, which deemed a transition non-modeled, if the monthly volume transitioning from delinquency bucket to the other were less than 10 loans /month. Given this logic, additional two transitions were determined ineligible for being modeled.
c) All remaining surviving transitions (those that survived Logic #s 1, 2 and 3) were tested based on Logic # 4.
d) Logic #4 deems a transition to be non-modeled if it fails the 1% contribution rule. Per this rule, for loans that transition from the source to destination bucket, if they represent either less than 1% of the total volume of the either the source or destination cell over the entire model development period, these are deemed ineligible to be modeled. Per this criterion, nine (9) transitions had failed and deemed ineligible for being modeled.
e) Lastly, for all remaining/surviving transitions (loans that survived Logics # 1-4), the C-statistics measure, as recommended by MRM's Model Testing Guidance as a key statistic to evaluate a parameter of importance, any transition that failed the statistical criterion of C-Statistic < 0.6[limit set within MRM's Model Testing Guidance] were deemed ineligible to be modeled as a logistic equation. Of all remaining transitions that survived Logics # 1-4, no transition has failed this criterion and hence were deemed eligible for being modeled as separate equations.
f) Summing up over all transitions that failed under each Logic of the waterfall, CAMU finally ended up with 28 equations that were deemed as non-modeled and the remaining two were deemed eligible to be – 'modeled' as separate equations. Please note there are seven equations in the entire RM transition matrix, which are considered as reference transitions.

g) Presented below is the logic used for estimating the non-modeled transitions. It is important to note here, that during the research and development of the logic that differentiated between modeled vs non-modeled transitions, several alternative approaches were also considered, of which the chosen approach was finalized due to simplifying the model complexity and at the same time effectively mitigating some of the prior model's limitations around recent period performances. All of the alternatives considered have been discussed in Section 3.2.2 of the MDD.

VS 3 Approach for non-modeled transitions

The non-modeled transitions utilized the following approach –

1. The Stress values are volume weighted mean of transition rates based on data from 200602 to 201112. The justification for doing so is that the US HPI reaches its peak in the second quarter of 2006 and started dropping over time. It reaches the bottom in 2009 and continues staying low until the first quarter of 2012. Instead of using the exact 27mo stress period defined by MRM, the stress value selection covered this entire period of HPI decrease to avoid over-fitting the model performance due to data volatility or at odd with MRM's "blind back test" requirement.

2. The non-stress/recent values are volume weighted mean of transition rates based on data from 201404 to 201712. The recent four years performance data well reflected the recent portfolio performance and go-forward portfolio mix while avoiding the data volatility and potential model over fitting using only the most recent two years' performance data. To be consistent with other model components, the same 201201-201403 performance data was hold out as OOT data and excluded from non-model transition assumption creation process.
3. Once the performance period is decided, CAMU considered segmentation impact to the non-modeled transition rates and created the stress/non-stress values separately for distinctive segments based on their empirical performance difference. There are three segments for RM included in trial loans, not in trial government loans and not in trial conventional loans. Delinquency information does not truly reflect borrower's payment behavior when a loan is in trial as the change of payment is not finalized and reflected in Bank's delinquency recording system. Therefore, abnormal delinquency movements are expected when a loan is in trial. CAMU has observed different non-modeled rates between government loans and conventional loans as well especially the increase of jump-to-worse transitions for government loans during the non-stress period. Separating government and conventional loans is a prudent choice to account for their difference over time.
4. Due to the uncertain and volatile nature of non-modeled transitions, the stress values are set to non-stress/recent values if the raw stress values are lower/higher than the non-stress values for to-worse/to-better transitions. This is a conservative treatment affecting limited cases (mostly observed for government loans jump-to-worse transitions) to ensure that the stress values are always worse than or at least the same as the non-stress values.

For more details on the methodology used for the calculation of roll rates for non-modelled transition for stress and non-stress periods, please refer to sas codes as mentioned within Section 5.1.9 of the

MDD. For additional details on why this approach was chosen as the preferred approach over other alternative approaches, please refer to Section 3.1 of the MDD.

VS 4 Approach for modeled transitions

After it was confirmed that there would be twenty-eight(28) transitions that would be modeled as equations in the 2019 RM PD model , the next step involved selecting the initial candidate pool(explanatory variables) for these twenty-eight(28) equations. The selection of all the explanatory variables for the initial candidate pool were based on the following criteria listed below:

1. Business Intuition / Economic Theory: Variables should make sense and there should be an economic rationale for their use in residential mortgage loan modeling.
2. Practical Aspects: Variables should (a) have sufficient historical data availability and (b) be readily available. (c) be well known. Forecasts may not be available for dynamic characteristic variables such as bankruptcy or automatic payment indicators.

All the variables used in the model development pertain to the following main categories-

- Loan Level Attributes-These describe the historical/current (marked-to-market) features of the loans such as delinquency history, Mark-to-market CTLV, Month on book, UPB, etc.) Most of the loan level attributes are dynamic in nature and updated periodically to reflect the most current specifics around the loan's performance
- Origination Profile-Origination features reflects the loan characteristics at the time of loan origination such as decision FICO, lien position, documentation type, origination channel, product type, loan term etc.
- Macro-economic Factors - such as HPI, Unemployment, interest rate, GDP etc.
- Bureau Attributes- (such as refreshed FICO, age of oldest mortgage, age since oldest trade, etc.

The variable selection for the 2019 Method A RM PD model started from the initial candidate pool of variables. The selection process was conducted separately for each of the transition equation in the model given that influence of loan characteristics, economic variables and payment patterns etc. will vary depending on the loan's current transition state. This differentiation helped improve the overall fit across transition equations and optimized the overall model performance.

After the initial compilation of the raw variables, the following steps were executed.

VS 4 Step 4.1

Conducted univariate, bivariate and correlation analyses on the overall pool of raw variables (inclusive of both numeric and character)

VS 4.1.a. Univariate analyses

Univariate analyses were conducted on all raw variables to understand the idiosyncrasies around the data fields. Univariate analysis results were used to ensure data consistency and completeness, given the modeling objective. The univariate analysis results were also used to understand data trends/outliers, missing values. Appropriate data treatments were meted out to the concerned variables, either in terms of missing value imputation (median +/- 2 standard deviation) and/or caps/floors. All necessary details around data treatment have been listed in Chapter 4 of the MDD

VS 4.1.b. Correlation

Correlation was conducted on all variables from step 4.1.b to streamline and reduce the number of candidate variables by selecting the top correlated variables by selecting the one with the highest correlation with dependent variable among the candidate variables with similar macroeconomic intuition. For example, when estimating the BKT01->BKT2 model, we use correlation analysis to determine that among 12-, or 6- or 3- month income change, only 12-month income change will be selected as the candidate variable as it's having the highest correlation with the dependent variable.

VS 4.1.c. Bivariate analyses

Bivariate analyses were then conducted on the remaining continuous variables after step 4.1.6 and corresponding bubble plots were created to understand the directional relationship between each candidate variable and the dependent variable and conducted suitable transformations (splines), as deemed necessary. Under the bubble plot analyses, first and foremost, insights are drawn based on shape of the curves (bubble plots) along with the size and spacing of the bubbles to incur if the relationship is in the expected direction. Second, it helped determine whether or not the empirical relationship is as strong as presumed based on common business intuition. Third, it provided a gauge of how significantly a change in this variable affected the probability of an event. A steeper curve implies a strong relationship while a flat bubble plot indicates a weak relationship between dependent and independent variables. Finally, the shape of a bubble plot demonstrates the extent to which the relationship is non-linear and prompted the modeler to apply appropriate transformations and/or create and justify usage of splines where deemed necessary.

While developing the 2019 Method A RM model, CAMU modeler conducted the above bivariate analysis for the portfolio segment such as by origination channel or by loan type (conventional and government loans). The purpose is to understand whether the binary relationship between the independent variables and the dependent variable varies across different segments. If needed, segment specific splines were created and tested in model estimation process.

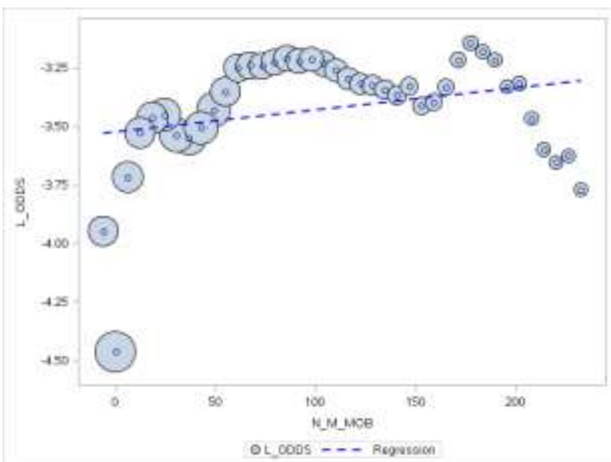
For example, figure 5.1.4.3 shows the bubble plot of MOB by loan origination channel created for the BUK01->BUK2 model. For loans originated via the correspondent, retail or broker channel, risk of delinquent increases with MOB after a loan was originated, and after the first 3 years, the performance gradually became stable and less sensitive to MOB.

On the other hand, the binary relationship was very different in the Wallstreet channel. Within the first three years, there was also an initial increase of risk with MOB for the Wallstreet channel. Compared to the loans with the same age but originated other channel, the Wallstreet channel has higher risk of delinquent. The observation is consistent with the general belief that loans originated via WallStreet channel were riskier because they were originated pre-crisis through sub-optimal underwriting standards.

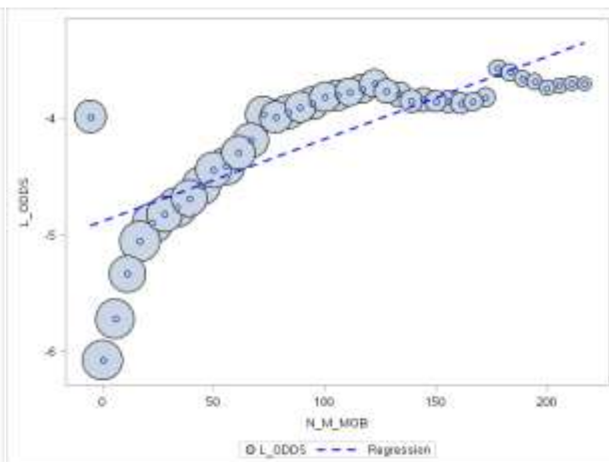
However, bubble plot based on the WallStreet loans also shows that their delinquency risk steadily decreases with MOB after the first three years. The loans that have gone through the macro economic stress period and keeps on staying at current bucket for a long time are more likely to continue to perform well. Risk difference due to loan origination channel became much less significant when MOB is long enough.

Figure 5.1.4.3 Bubble plot of MOB effect for BUK01->BUK2 Model by Origination channel

Correspondent Channel

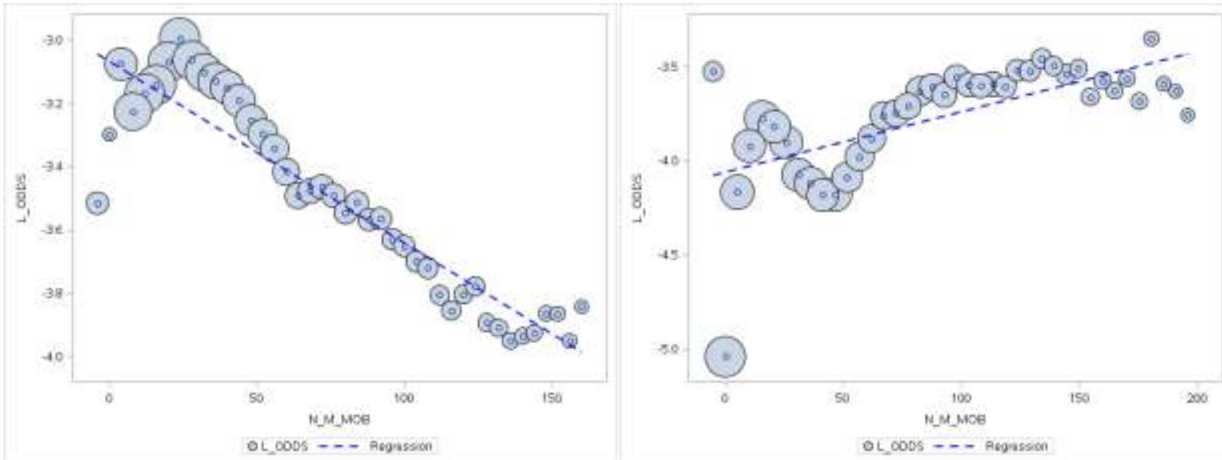


Retail Channel



Wallstreet Channel

Broker Channel



Based on the bivariate bubble plot, CAMU determined it's prudent to create and test a MOB spline at 36 months, along with creating the interaction term between MOB and Channel to test the MOB effect specifically for each channel.

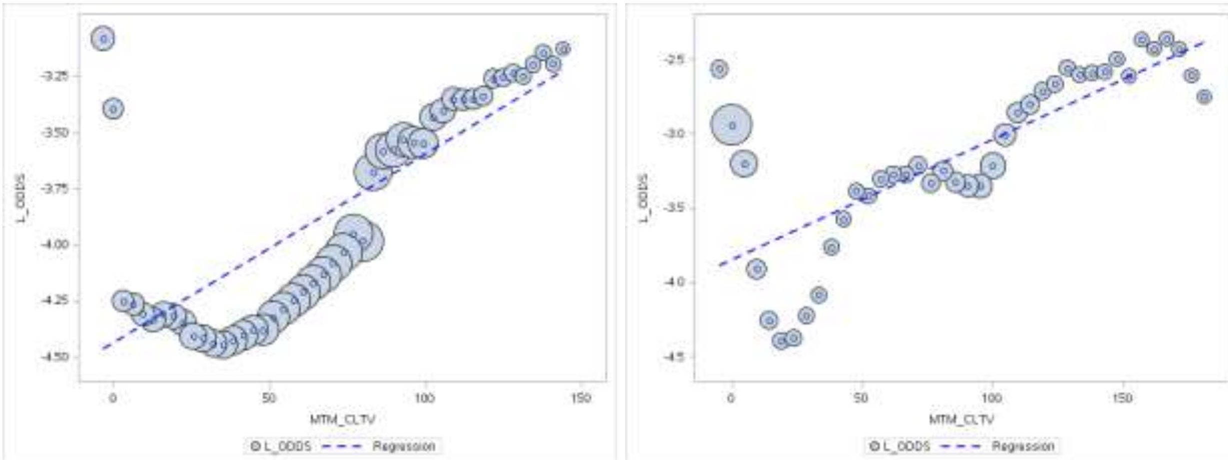
Another example is the Mark-to-Market CLTV (MTM_CLTV) effect in the BUK01-BUK2 model, specifically conducted on conventional first lien, conventional second lien and the government loans. As show in figure 5.1.4.4, for conventional first lien, the delinquency risk increases with MTM_CLTV, so the risk of a loan with 40% CLTV is significantly lower than an 80% CLTV loan. Meanwhile, the risk sensitivity to CLTV became less strong when MTM_CLTV is greater than 80. However, a more non-monotonic relationship was observed on the second liens. There is almost negligible differentiation when CLTV varies between 40% and 80%.

Based on the bubble plot, CAMU created CLTV splines at 40% and 80%. The two splines were interacted with conventional first lien dummy and conventional 2^{nd} lien dummy to specifically capture the different CLTV effect on these two group of loans.

The binary relationship between CLTV and delinquency risk is very different on the Government loans. The initial observation that is consistent with business intuition is those loans whose CLTV is greater than 100% is significantly higher than those whose CLTV is lower than 100%. CAMU further found that the extremely low risk of Government loan at around CLTV equals to 100% were loans that just originated, namely, the ones with N_M_MOB <=3. CAMU then split the GOV population into two groups, N_M_MOB<=3 or N_M_MOB>3. Within these two groups, the bubble plots shows an intuitive relationship where the higher the MTM_CLTV, the more likely a loan will go to delinquent. So CAMU created a dummy indicator for GOV loans with MOB>3, and interact with MTM CLTV to capture the government loan sensitivity to MTM CLTV. The interaction term was tested and kept in the final BUK01->BUK2 specification.

Figure 5.1.4.4 Bubble plot of MTM CLTV effect for BUK01->BUK2 Model by Loan Type and Lien Position

Conventional first lien | Conventional second lien



Government Loan

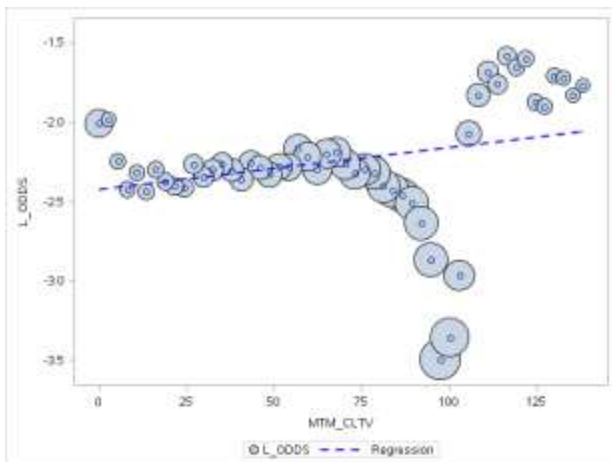


Figure 5.1.4.5 Bubble plot of GOV MOB effect in BUK01->BUK2 Model

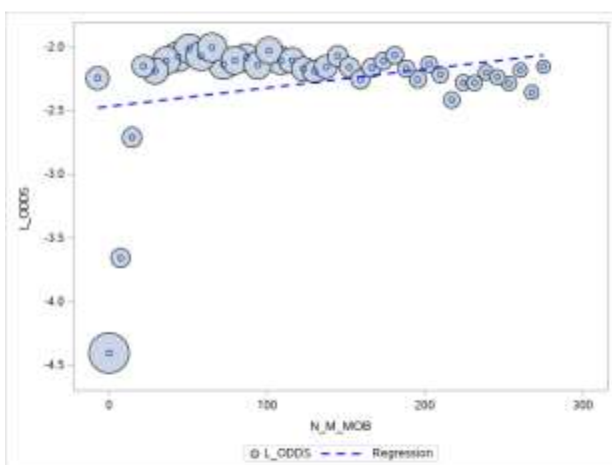
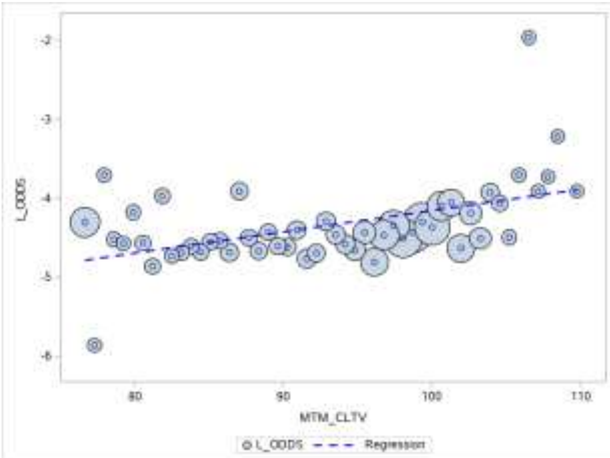
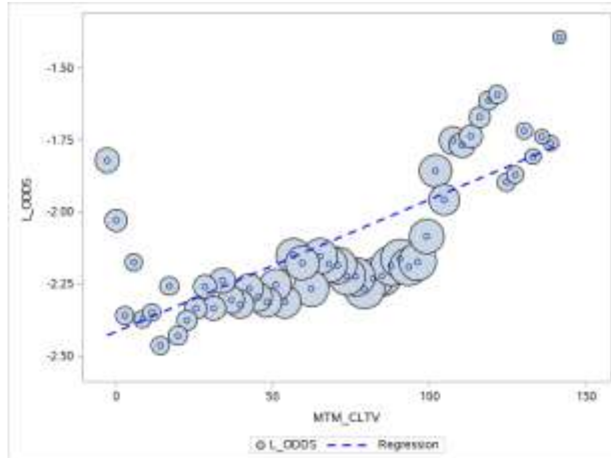


Figure 5.1.4.6 Bubble plot of GOV MTM CLTV effect in BUK01->BUK2 Model

GOV loan with MOB<=3 GOV loan with MOB>3





VS 4 Step 4.2

After the completion of all univariate, correlation and bivariate analyses, the next step involved executing regression analyses in an iterative process using the backward selection method to manually select/drop variables from the modeled equation(s). To simplify, all raw variables were initially added to the regression model and consequently predictor variables were removed in a recursive process based on the significance level, in a stepwise manner until there is no variable left to be removed anymore and a parsimonious model is obtained. To provide some additional context, on the fine-tuning of the modeled equations through its parametric specifications, it is important to get some context on the overall Method A modeling framework. The Method A PD model follows a transition framework wherein the state of the current period is dependent on the state of the prior period. Under this framework, the dependent variable is the migration or transition rate from a source delinquency bucket to a delinquency bucket of choice rather than to migrate into the reference delinquency bucket which is influenced by the borrower credit characteristics, macroeconomic factors, past delinquency status/payment pattern and other origination/collateral specific features. The complete m x n transition matrix within the Method A PD model represents the month-by-month movement of loans between current, delinquency, VP, and IVP states. Each of these transitions is separately handled/modeled within the Method A PD model suite, based on their unique risk drivers. As mentioned earlier, the inclusion/exclusion of attributes within each transition equation has been achieved using backward selection. The Selection Process around the independent variables has been summarized below:

1. Tested all candidate variables that were selected after the correlation analysis in step 4.1.c, including their transformation or splines, on the development sample using logistic regression
2. Carefully reviewed the regression output and excluded variables that were statistically significant (p-value <=0.05) but had counter-intuitive effects.
3. Re-estimated the logistic regression based on the remaining candidate variables.

4. Removed variables that were not statistically significant, starting from the ones with the smallest chi-square, and repeated step 3 iteratively. Please note this modeling exercise never excluded all insignificant variables at one time. Instead, this step was repeated several times and in each run, only 3 to 5 least important variables were dropped. Once an effect was removed from the model, it remained excluded. The process was repeated until no other effect in the model met the specified level for removal.

5. Variables that breached the variance inflation factor (VIF) > 5 threshold in the collinearity test (VIF test) were excluded. If variable A and B are found to be highly correlated with each other with both demonstrating high VIF, the modeling team modeler tested the inclusion of only one of them in two independent runs and kept the one that exhibited a better fit. Note at this step, exceptions were made if the splines and their raw variables, such as MTM_CLTV and MTMT_CLTV_SP80, co-existed within the same model.

6. Tested the selected variables on the 20% in-time hold-out-validation sample data and full sample (which include both in time and out-of-time sample). Variables that demonstrated opposite effect or became statistically insignificant or experienced large change in coefficients in different samples were excluded.

7. The Akaike information criterion (AIC) was checked to obtain the best trade-off between model fit and parsimony. In particular for the independent variables with low chi-square, additional test was conducted to see if their removal led to a lower AIC. If it did, these variables were excluded.

To summarize, the following criteria was adhered to justify the overall variable exclusion process, in a iterative fashion-

To remove/add a variable from a given equation, the following rules are adhered to-

- Variable significance and parameter sign
- Multi-collinearity (VIF) < 5
- Consistent with business interpretation/economic intuition
- Robust in terms of model's goodness of fit across in-sample and out-of-sample based on model's In-sample goodness-of-fit
- Parsimonious with reasonable AIC statistics and one-month ahead fit

VS 4 Step 4.3

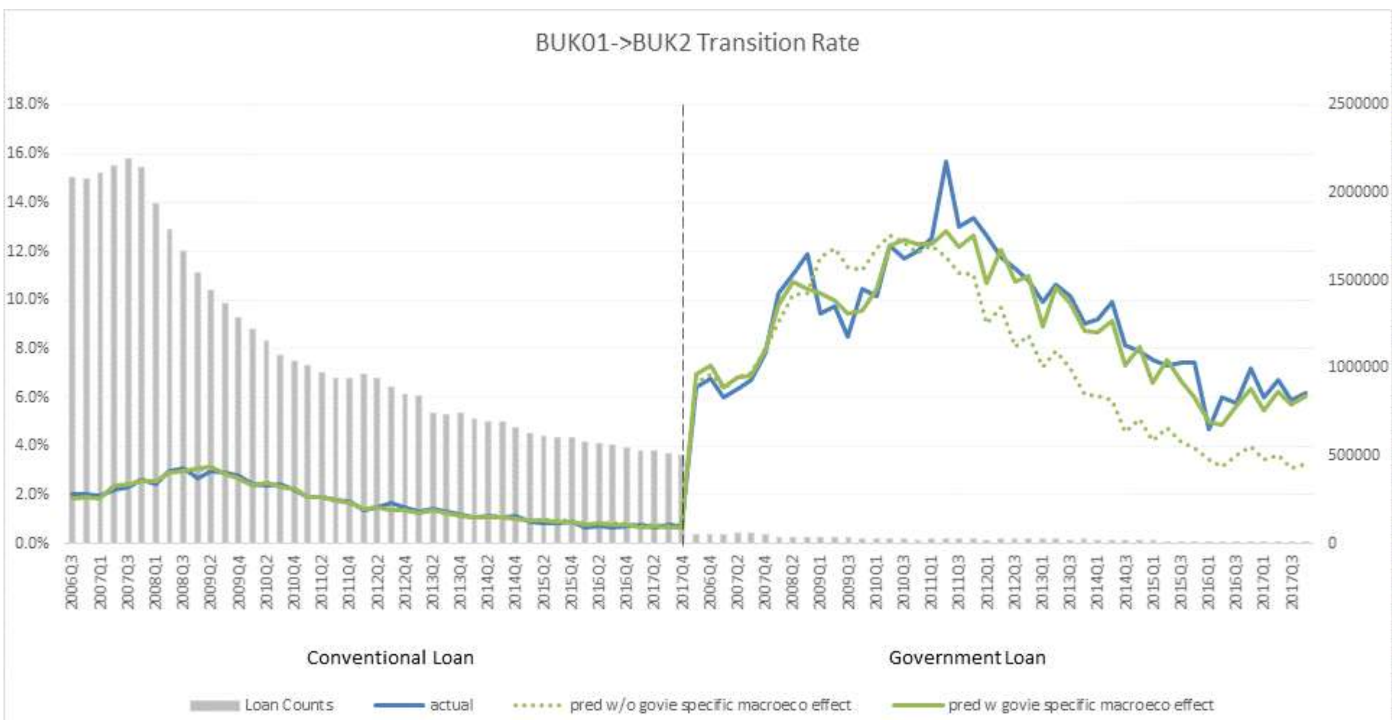
Once the preliminary model specification was built based on both qualitative and quantitative evidence, the next step involved analyzing the modeled equation's one-month ahead performance(actual vs predicted) by varied key risk drivers/segments(like origination channel, FICO, LTV, etc.) across both stress and recent periods. The fundamental concept underlying this deep-dive analysis was to make sure that the modeled transition performed well in each sub-segment of a given key risk driver. If any systematic under/over prediction were observed for a given sub-segment for a given risk driver, appropriate dummy effects were introduced in the model to control for this effect. Similarly if certain sub-segments demonstrated differentiated risk sensitivities to macro-economic indicators compared to the given norm, suitable interaction variables were added and tested in the model to capture the effect of these differentiated sensitivities. The addition of interaction variables had a two-fold advantage. In essence, they were able to capture the disparate effects of the macro-economic indicators on certain sub-segments of the modeled equation, without having to introduce a separate equation for this sub-segment, thus preserving the model's simplicity.

Still use the BUK01->BUK2 model development as the example. As shown in figure 5.1.4.5, as represented by the grey bar chart, conventional loans took up the overwhelmingly larger portion in the development data than the government loans. Based on the blue line, for the conventional loans, the actual peak delinquency rate happened in late 2008 and early 2009, when the year over year HPI deterioration and the unemployment rate were both at the highest level. The peak delinquency rate of the government loans, however, happened in early 2011, indicating that the government loans may have a delay response to the macroeconomic risk drivers.

Initially the model was estimated without considering government loans specific macroeconomic effect and without any lagging effect to the macro economic variables such as HPI change or unemployment. When examining the one-month-ahead forecasting accuracy, CAMU modeler noticed that the predicted delinquency rate of the government loans, as denoted by the green dashed line, had a very similar historical pattern as the conventional loans. The model systematically over-forecasted the government loan BUK01->BUK2 transition rate between 2009 and 2010, but under-forecasted after 2010.

Model developer considered the lagged macroeconomic effect, such as year over year HPI change 12 months ago, and unemployment rate 12 month ago, into the model. CAMU then created and tested interaction variables between the macro economic variables and conventional/government loans indicators. As shown in the green solid curve, the model with separate macroeconomic effect on conventional/government loans is able to capture historical BUK01->BUK2 roll rate accurately for both conventional loans and government loans. In the final model specification, government loan transition rate is significantly positively associated with the state-level unemployment rate 12 month ago.

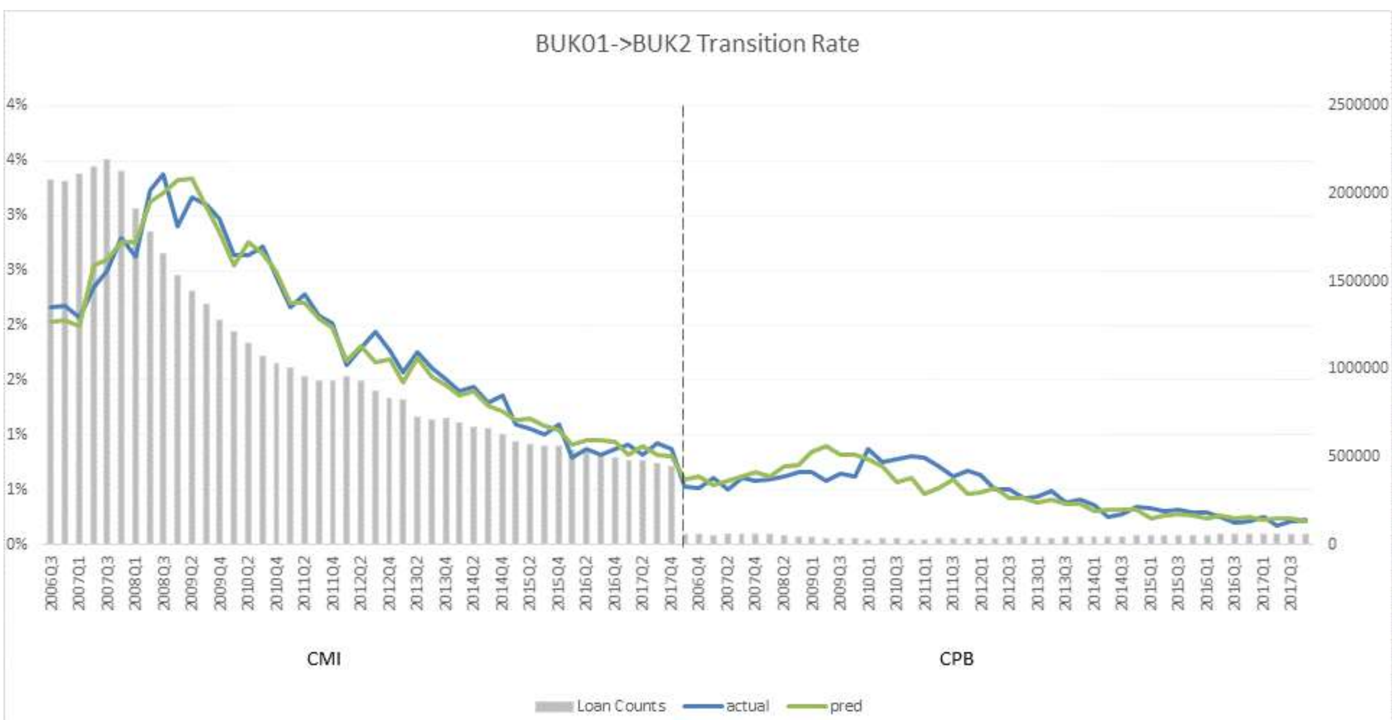
Figure 5.1.4.5 Actual and predicted BUK01->BUK2 transition rate by loan type



Similarly, by checking the one-month-ahead fit when separating the portfolio into CMI and CPB, CAMU also noticed that the model, which was predominated by the performance of the CMI portfolio, forecasted a higher than actual roll rate on the CPB portfolio in 2009, but an lower roll rate in 2011. Figure 4.1.4 below also suggested that a lagged macroeconomic effect should also be tested specially for CPB. So CAMU also carefully tested variables such as year over year HPI change 12-months ago, and unemployment rate 12 month ago, and 2-year HPI change rate on only the CPB loans. However, none of these variables could be brought into the model because the signs associated with them were counter-intuitive in regression.

An important point to note here is that while the one-month ahead analysis did reveal that some transitions shared similar patterns, it was considered prudent to not combine these transitions together. In fact, to add some context this was one of the alternative approaches that was tested during the initial model development phase. For additional details on all approaches that were tested, please refer to Section 3.2.2. of the MDD. Under this approach, the plan was to combine transitions which exhibited similar sensitivities to key risk drivers & to macroeconomic variables with the addition of appropriate dummy effects to capture the nuances of each specific transition. Please note that this approach had been used by the Method B model suite (2017 Challenger model) for the 2018 CCAR process. However when it came around to evaluating the effectivity of this approach using model's backtesting, it was noted that the backtesting itself required the re-distribution of the collated loans which became a modeling challenge, given the limited volume associated with most of these transitions. Incorrect loan allocation during backtesting would have affected the marginal predictions of the loans in the specific buckets which could have potentially impacted model's performance. The same concern was also raised by the model reviewers' on the Method B model suite during the 2018 Model Validation Exercise and CAMU used the past experience as lesson learnt and considered appropriate to drop this approach from further consideration.

Figure 5.1.4.6 Actual and predicted BUK01->BUK2 transition rate by business line



VS 5 Finalization of Specifications

Once the parametric specification of all transitions were finalized, the modeled equations were further fine-tuned and modified based on the following conditions-

- Model specification comprehensive review (internal to CAMU)
- Back test analysis, Characteristics analysis
- Sensitivity analysis

The final specification of the Method A PD model is provided in attachment '5.1.4 Residential Mortgage Model Specification.xlsx'. The justification for every variable in the final model specifications is discussed in more detail in MDD section 5.2 below.

To summarize all PD model equations have been validated across time, based on all the criteria listed above for both in-time and OOT validation time frames. All the key risk drivers in the PD model remain significant in the 27 month forecast period as validated within the model's robustness and stability test results as discussed in section 6.2.

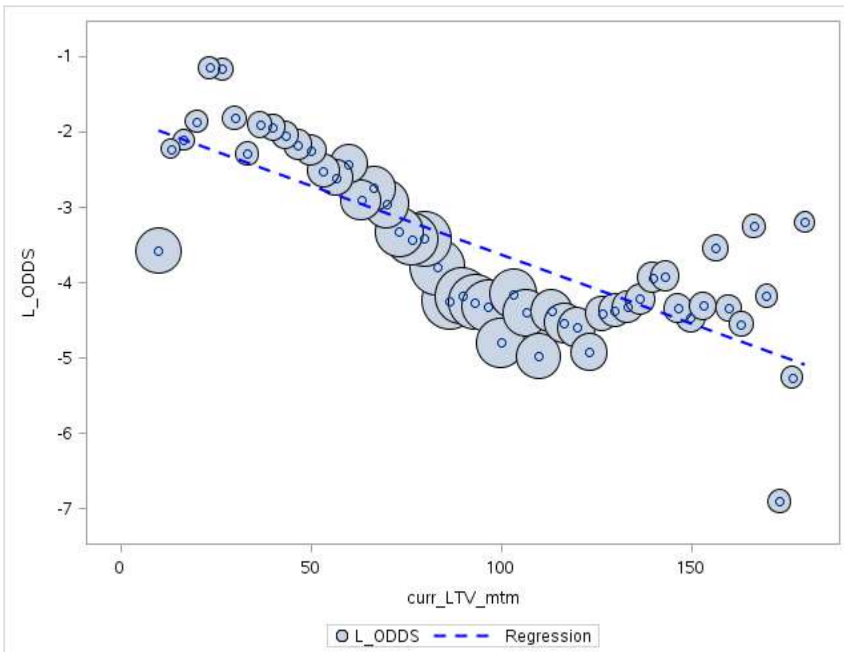
LGD Model Variable Selection

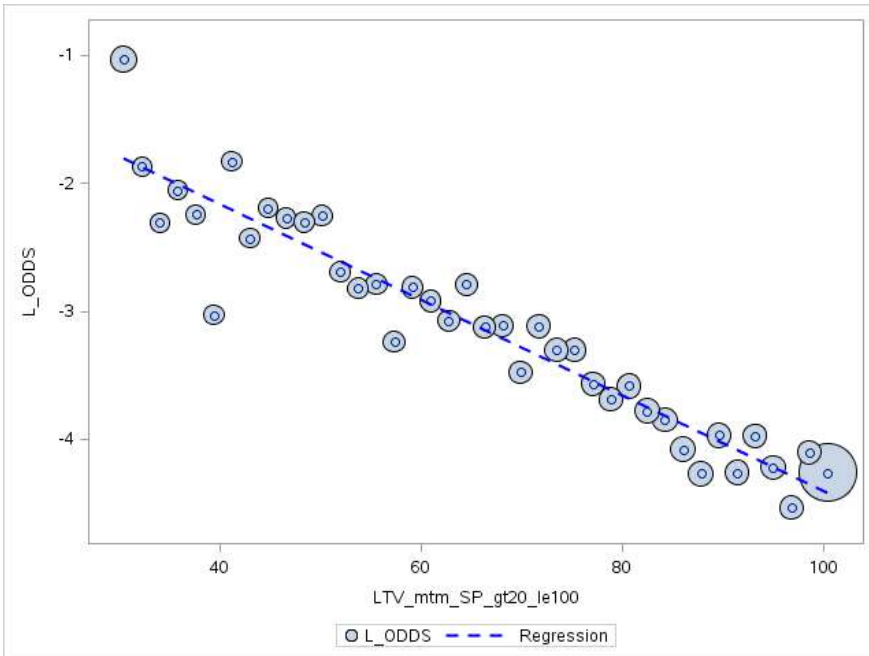
The variable selection procedure for the Method A severity model started afresh with considering all variables that have been deemed important for the LGD model. In the process of creating the initial candidate pool of variables, considerations were given to the prior year's candidate pool of variables and the reviewers' feedback on the overall variable selection process. Additionally, based on the targeted improvements attempted this year along with the new modeling methodology, new attributes were added to improve model's performance across chosen segments of interest.

The LGD variable selection followed the same procedure as PD, essentially for each explanatory variable: their intuitive meanings, parameter sign direction, significance (p-value), VIF, robustness (long model) were checked. As part of the process, some variables were removed either due to counter intuitive interpretation or due to not being statistically robust. Additionally, the modelers also examined the completeness of the input information (predictors), evaluated the LGD model's fitting in-sample (200801- 201112, 201404-201706) and OOT (201201-201403) samples, as well as its back testing performance. The LGD model specification and performance were extensively reviewed and variables' justification challenged as part of internal model review sessions. All feedback was captured and was incorporated in the model development process to further fine-tune the model and improve the model's performance. The LGD results for the first and second lien models finally yielded the model specification as cited and discussed in Section 5.1.4. Please note, given the two stage modeling approach for the LGD model, the model specifications are different for the first stage model which models the propensity of incurring zero/full losses relative to partial losses. The second stage of the model then estimates the partial loss rate (zero and full losses are assigned a 'zero' and 'one' probability respectively). The Full loss are defined as Loss_rate>0.97 while the zero loss is defined as Loss_rate<=0 and partial losses as 0<loss_rate<=0.97, respectively.

In the finalized LGD model, both univariate and bivariate analyses were conducted to gain understanding on the predictor variables and improve the model's fit accordingly. Based on the univariate analysis, suitable data treatments were carried out, all of which have been discussed on chapter 4 of the MDD. Based on the bivariate (bubble plots) analyses, variables were transformed or splines introduced as deemed necessary. For example, in the first lien stage 1 – Zero vs Partial loss model, marked-to-market LTV was floored at 20 and capped at 100(LTV_mtm_SP_gt20_le100) to maintain the model's monotonicity to increasing LTV.

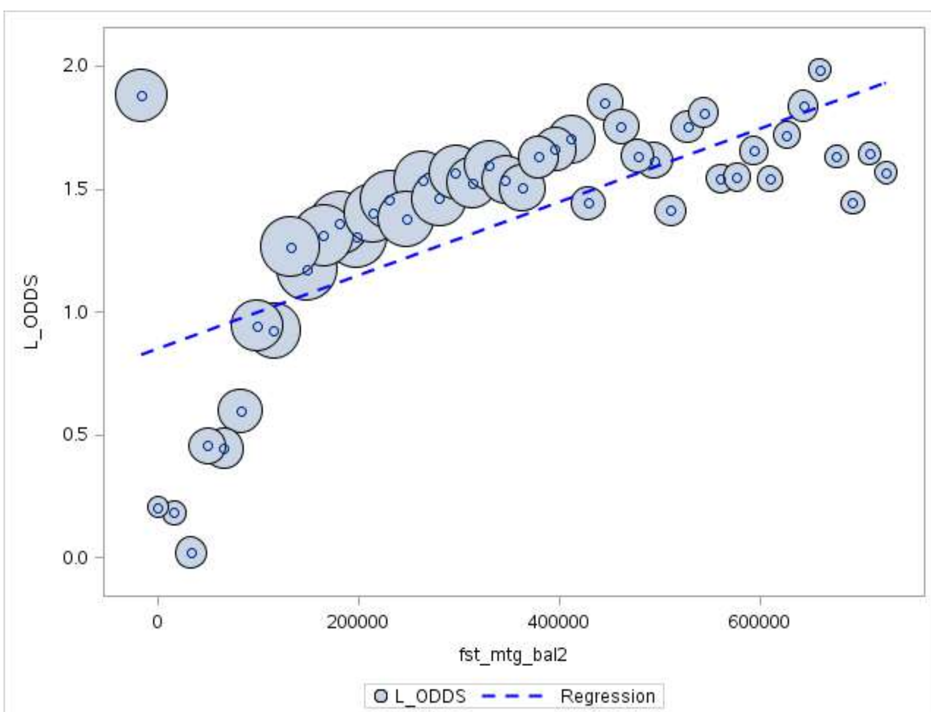
Figure 5.1.4.7: Spline Transformation for Marked-to-Market LTV

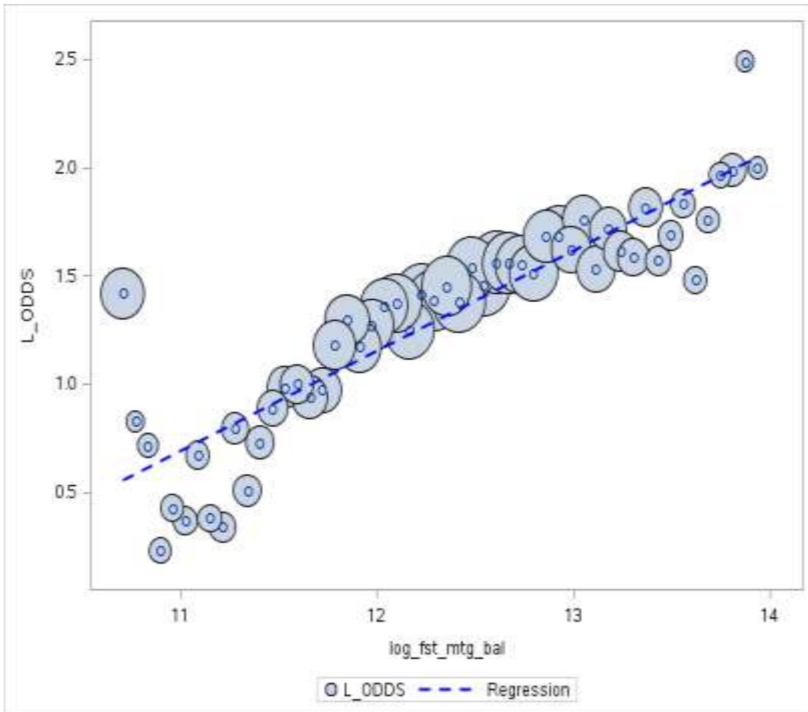




Log transformation also used to improve linearity between dependent and independent variable. For example, in second lien stage 2 model, log transformation of first lien mortgage balance is used, as illustrated below.

Figure 5.1.4.8: Spline Transformation for First Mortgage Balance





LGD for CPB Loans

As described in Chapter 3 – Modeling Approach/Methodology, the CPB portfolio is characterized with a sparse loss population. To provide some context, the entire development data witnessed only 300 accounts that encountered losses, out of which approx~ 260 accounts pertained to the first liens and the rest ~50 accounts corresponded to the second liens. Since the limited data created challenges in building an economically meaningful credit loss model, with extremely low volume for the second lien loans, it was considered that all these ~50 loans would be assigned full losses. For the first lien loans, the base losses were based off an empirical lookup table that was created using historical CPB loss data. For the stress scenario, the modelers estimated a scalar based on the stress to base ratio for the first lien loans off the CMI portfolio. The CMI portfolio was chosen to calculate the scalar as it had a rich history of data to effectively calculate the stress to base loss differential. The first lien base case losses were adjusted by multiplying with this derived scalar to get to the stress losses.

For further details of the 2018 first Mortgage LGD variable selection process, including explanatory variables and final model parameter specifications, please see the attachment 5.1.4 Severity_model_specifications'.xlsx.

If multiple dependent variables were tested during development of the model, why was the specific variable used in the final model chosen over the other alternatives. If transformation is used provide rationale behind the selection of the chosen transformation and its appropriateness to the model objective? Refer Model Testing Guidance with regards to Dependent Variable construct

For Residential Mortgage PD model, each modeled equation has been developed with one dependent variable, although the transition system of equations has multiple outputs including delinquency status, involuntary payoff (IVP) and voluntary payoff (VP). For LGD model which follows a two-stage approach to modeling losses, wherein the first step is a binary equation that models the propensity of full/zero losses relative to partial losses and the next stage models the partial losses.

Is the variable selection technique applied robust over time (i.e. is it invariant to the starting point)?

Yes. The variable selection procedures have been carried out with a development dataset which includes a full business cycle with both stress and non-stress periods. Further, the one-month-ahead analysis and back testing over different periods ensured that the variable selection process was robust. As part of variable selection process, for both PD and LGD models CAMU built the corresponding equations using development, in-time hold out validation and full sample. The parameters estimated on the development and validation samples. The parameters estimated on the development and validation samples were compared to make sure both PD model and LGD model are robust. Refer to MDD Chapter 6 for all testing details.

What is the rationale for applying a particular selection technique? For example, why a VarClus procedure was chosen as the variable selection technique over others?

As described above, no single variable selection technique is used. Rather, univariate, bivariate and correlation analyses techniques were deployed and the final package is based on multivariate techniques such as one-month-ahead accuracy and backtesting accuracy, which uses the entire system of equations.

Do the independent variables make sense for use in the portfolio? In particular, was appropriate justification provided to demonstrate that all relevant variables were incorporated into the model? Was the business rationale behind all variables' relationships with historical data clearly explained and appropriately justified? (i.e., all relevant drivers are incorporated into the model and there are no variables included that have a spurious relationship with the modeled quantity)

As widely known, credit risk management and the adequate modeling of credit risk exposure for consumer portfolios have received tremendous attention from both the regulators and the banking industry in recent times. The credit quality of a given loan at a given snapshot and the dynamic transition of the loan into subsequent worse/better delinquency statuses, and into prepayment or default are strongly influenced through systematic (macro-economic, current lending practices, credit policy, etc.) and idiosyncratic (loan and borrower specific) factors.

The 2019 CCAR Method A model redevelopment exercise considered an exhaustive list of key risk drivers spanning across macro-economic attributes, loan origination and dynamic features, borrower attributes and past payment/delinquency history and natural disaster attributes as part of its variable selection process, for the estimation and prediction of key risk modeling parameters like Probability of Default (PD) and Loss Given Default (LGD). Based on model sponsor leadership requests, effects from natural disasters such as hurricanes were included in the list of risk drivers considered in the 2019 model redevelopment project in the form of indicator variables. Both PD and LGD are critical model inputs which ultimately feed in the estimation and forecasting of gross/net credit loss exposures for the North America Real Estate portfolios. These loss estimates, apart from fulfilling CCAR Reporting requirements, also help drive Risk Appetite Framework (RAF), Loan Loss Reserving (LLR) Troubled Debt Financing (TDR) processes. A detailed description of the model specific usages/applicability has been outlined in Chapter 1 of the MDD.

As already stated, all of the independent variables were selected based on the combination of business practice, statistical analysis, literature review, and industry approach. The list of risk drivers includes both macroeconomic variables and loan specific characteristics.

It's important to note here that, each variable included in each model has been thoroughly reviewed and challenged by CAMU peer modelers and the model's functional reviewers. The independent variables' relationship to the dependent variable in historical data have been clearly explained and appropriately justified by appropriate business rationales to establish adequate reasonable basis around the selection of a given variable in the final model for forecasting purposes.

Note – Additive Effect for some attributes:

CAMU would like to point out that for both the PD and LGD model there are NO inconsistent or non-monotonic parameter sign issues. CAMU urge model reviewers' to carefully note the variable explanations/definitions column while reviewing model specifications as some attributes and their splines may have an additive effect. To give an example, let's say a modeled equation possesses both the original attribute (like MTM_CLTV) and its splines for less than 40, 40 to 60, 60 to 80 and 80 and above.

The interpretation of the parameter estimates for these attributes should go like this -

MTM CLTV – main MTM CTLV effect;

MTM CLTV for spline less than 40 – Parameter sign and estimate is an additive effect of MTM CLTV and MTM CLTV for spline less than 40;

MTM CLTV spline between 40-60 - Parameter sign and estimate is an additive effect of MTM CLTV and MTM CLTV for spline less than 40 and MTM CLTV spline between 40-60;

MTM CLTV spline between 60-80 - Parameter sign and estimate is an additive effect of MTM CLTV and MTM CLTV for spline less than 40 and MTM CLTV spline between 40-60 and MTM CLTV spline between 60-80;

MTM CLTV spline greater than 80 - Parameter sign and estimate is an additive effect of MTM CLTV and MTM CLTV for spline less than 40 and MTM CLTV spline between 40-60 and MTM CLTV spline between 60-80 and MTM CLTV spline greater than 80;

PD Model

The variables considered as part of the Probability of Default (PD) model development exercise have been classified in the following categories, segregated based on their destination status migrations which can be either terminal (voluntary prepay -VP, involuntary prepay-IVP) states and/or to worse and to better (cure and partial cure) delinquency states. The main rationale behind this segregation is that for all equations that are moving in a similar destination (VP or IVP) or a similar direction (to-worse or to-better) should exhibit similar economic intuition and consistent parameter sign. So for example, if we consider MTM CLTV, customers with higher CLTV would more likely to go to worse, keeping everything else constant. So for all delinquency statuses (or buckets) which transition to deeper delinquent bucket as their destination cell, CLTV as the predictor should have same interpretation with a positive correlation. Presented below are the variables and their supporting rationale, for each destination status.

A) Voluntary Payoff (VP) transitions

Voluntary Payoff (VP) is the early prepayment of a loan by a borrower, in part or in full, often because of optional refinancing to take advantage of lower interest rates. The following attributes came up as significant predictors of the residential mortgage VP movements across all transitions.

Present Interest Rate Spread – The present interest rate spread is the difference between the current note rate that a borrow has and the current market rate. Lower spreads are indicative of economic well-being with borrowers being perceived to have lower credit risk and vice versa. For a FRM loan or pre-reset ARM loan of which note rate is fixed (i.e., note rate does not change along with the current market mortgage rate), the higher spread yields higher opportunity to refinance at lower rates. In addition, prepayment sensitivity to the spread varies across portfolio (CMI vs. CPB) and vintages (pre-2010 vs post-2010). Several interest rate spread variables were included in the VP model as interaction attributes. Please refer to the specifications file for additional details. Please note that the interest spread only appears in the BUK01->VP transition. The reason why this attribute does not appear in the other VP equations is primarily due to the fact that they appear insignificant in the model and were dropped as part of the comprehensive and iterative variable selection process.

Marked-to-market CLTV and its splines- A marked to market CLTV is determined by adding the balances of all outstanding loans and dividing by the current market value of the property. A marked to market CLTV provides a more accurate representation of a loan's credit risk that is encumbered by all liens, compared to the original CLTV. Higher values of the marked-to-market CLTV imply lower equity in the property, and therefore lower probability of prepayment. The marked-to-market CLTV has appeared in all VP transition equations as one of the important predictors. Additionally, VP also exhibits different sensitivity to the marked to market CLTV across different portfolio (CMI vs. CPB) and under different economic conditions. For example, the marked to market CLTV is also interacted with periods of decreasing HPI to capture the idiosyncratic behavior of borrowers (essentially borrowers lack the willingness to pay as their confidence in the state of economy decreases with deteriorating HPI) with MTM CLTV > 60 during periods of depreciating housing prices. The CPB borrowers are less sensitive to changes in macroeconomic attributes and do not necessarily exhibit any sensitivity to prepay even when the equity in their home increases. Splines are used to capture the differentiated sensitivities at different slopes (floor at60, floor at 60 and cap at 80 and cap at 100). For justification around spline for BUK1->VP, please refer to attachment -5.1.4 Bubble Plot Justification.xlsx. For details, please refer to 5.1.4 Residential Mortgage PD Model Specification.xlsx.

State Level Unemployment - It measures the prevalence of unemployment in the state and is calculated as a percentage of the number of unemployed individuals by all individuals currently in the labor force in that state. During periods of recession, an economy usually experiences a relatively high unemployment rate. High unemployment rate and (or) unemployment increase suggests higher probability of default and lower probability of prepayment. Both unemployment rate and its rate change (12m) are included in the VP equations. The unemployment rate is used in all VP equations to demonstrate lowering of the borrower's ability to pay when unemployment increases. The short-term unemployment rate change reflects the additional constraint on the already delinquent borrowers who have additional difficulties finding alternative sources of income. For details please refer to 5.1.4 Residential Mortgage PD Model Specification.xlsx.

FICO- FICO scores stem from modeling pioneered by Fair, Isaac and Company (now known as Fair Isaac Corporation). A borrower's FICO score is deemed to be an important predictor for probability of prepayment since it measures borrower's credit quality to refinancing elsewhere. A high FICO implies better credit quality; hence high probability of prepayment, ceteris paribus. For the 2019 CCAR model development exercise, the FICO score came up as one of the key prepayment-modeling attribute. Both Decision FICO and Refresh FICO are used in the VP equations. In fact, as can be seen the Refresh FICO is only used in the BK1->VP equation. The rationale behind not introducing Refresh FICO in any other delinquent bucket is because the model doesn't forecast refresh FICO, as doing so can lead to unstable model errors associated with using the pre-populated Refresh FICO as of the snapshot month. So for all delinquent buckets > 1, the usual practice is to leverage the decision fico.

Number of months where market rate is lower than origination rate- It is used to capture the burn-out effect of prepayment. As the mortgage rates continue to fall the prepayments will begin to level off as only those borrowers who are the most refinancing impaired or who do not have willingness to refinance remain in the pool, this is known as the burn out effect. Typically, a loan that is experiencing lower market rate than its origination rate is considered to have incentives to refinance. Therefore, the higher the number of months that a loan has experienced rate lower than origination, the lower the expectancy to voluntary prepay.

MOB and its splines: This is a numerical indicator for how long a given loan stays on Bank's portfolios. The month on book variable has been divided into spline transformations to better capture the nuances of the effect of loan age on its credit performance. It has been observed and documented in literature that, all performing loans (not delinquent) within 36 months on book, are more likely to payoff indicating the fact that unseasoned borrowers are more likely to actively shop for lower rates or sell the property within a span of three years. For loans that stay on book, beyond 36 months, are less likely to payoff, indicating their seasoned nature. For additional justification, please refer to attachment - 5.1.4 Bubble Plot Justification.xlsx

Loan-in-Trial Indicator- For loans that qualify for a loan modification, there is a precursory trial period, wherein the lender facilitates the paperwork for the loan modification process. Typically the length of the trial period varies based on the type of the loan and the lender requirements. When loans enter the trial period, it indicates both the borrower and the lender are actively working towards modifying the terms and conditions of the loan, including but not limited to lowering the interest rate, reducing the number of payments or principal balance. All loans that are in the trial period are less likely to prepay as they are actively locked down in the trail modification process.

FRM Loan Indicator- A fixed-rate mortgage charges a set rate of interest that does not change throughout the life of the loan. Although the amount of principal and interest paid each month varies from payment to payment, the total payment remains the same. In the 2019 Method A Residential Mortgage PD model, FRM with a term greater than 15 years appears in almost all of the VP transition equations and exhibit a negative correlation with the dependent attribute. On one hand, the 15-year loan is more likely to have ability to refinance to a shorter term product like 10-year loan which has significant lower interest margin. On the other hand, people originating a 15-year loan are considered to be more favorable to low interest rate and thus would be more sensitive to rate changes and more active in refinancing than people originating a 30-year loan. Another attribute of somewhat lower interest is the interest rate spread for a FRM loan. Typically the greater the spread, the greater the incentive to refinance a loan, hence interest rate spreads for a FRM loan are positively correlated with the VP parameter.

Reset Indicator (ARM/IO) – A mortgage reset is the point in time at which the mortgage rate and payment changes. It typically happens for ARM or interest only (IO) loans. ARMs have a fixed period of time during which the initial interest rate remains constant, after which the interest rate adjusts at a pre-arranged frequency corresponding to the changes in the indexed rate (Prime Rate, Libor, Treasury note, etc.). An interest only loan is a loan where the borrower only pays the interest on the mortgage through monthly payments for a term and then after this term the borrower start to pay both interest and principal balance Given the increase in rates for ARM or payment increases for IO, it is not uncommon for borrowers to refinance to lower rates once the loan hits the reset stage, to keep their monthly payments at a lower level. As expected, loans that approach the reset (defined as pre-reset 6 months or 12 months) or have reset in the recent past (defined as past 3 months of reset for IO loans and 12 months post reset for ARM loans) are more likely to refinance to lower rates and hence are positively correlated with voluntary prepay.

ARM Indicator (greater than 7 years) - ARMs or Adjustable Rate Mortgage**s** usually feature lower rates and payments early on in the loan term and allows borrowers to take advantage of falling rates without the need to refinance. This in turn provides a cheaper alternative, especially for borrowers who do not plan to live in one place for a long time. Typically ARMS which have the introductory rate(lower payment) lasting more than 7 years are less likely to prepay as these owners are less likely to flip over their property compared to owners whose ARMs have a shorter teaser rate periods(3/1ARM, 5/1ARM). Owners with ARMs less than years are investor – buyers who aim to buy and retain the property for a short period and then sell it, making profits on capital gains.

Community Loan Indicator- Community loans fall under the category of special lending programs often leveraged to provide financing to low- and moderate-income borrowers. Bank's administers these programs to prudently serve the credit needs of the bank's communities, with the additional benefit of potential consideration in the assessment of the bank's performance under the Community Reinvestment Act (CRA). While the management has set clear risk-adjusted policy limits for commitments to these programs, it is important to note that the Community-loan home-buyer programs with limited or no down-payment requirements may permit borrowers to become overextended on their credit, more easily than conventional mortgage programs. Overextension is also more likely to add another layer of credit risk through increased LTV ratios, higher DTI ratios, or reduced savings on hand. Any or a given mix of these credit specific attribute(s) of layering risk can result in a borrower quickly becoming delinquent/default due to an unexpected financial hardship triggered either by the general macro-economic outlook or a specific change in borrower's financial health. All community loans, given their origination specific characteristics are less likely to VP (prepay/refinance)

Principal Balance and its spline- This attribute is expected to be positively correlated with VP transition with the rational that higher balance loans have higher prepayment incentives. By using a spline transformation (spline flooring at 50K and capping at 100k) , this positive effect starts to be effective from $50,000 and becomes a flat effect after the balance reaches $150,000. It is consistently used in current and low delinquent buckets to VP models. For added justification on the spline usage, please refer to attachment 5.1.4 Bubble Plot Justification.xlsx

Prepayment Penalty - A prepayment penalty is a clause in a mortgage contract stating that a penalty will be assessed if the mortgage is prepaid within a certain time period. The penalty is based on a percentage of the remaining mortgage balance or a certain number of months' worth of interest. Loans with any sort of prepayment penalty typically are less likely to prepay as rates go down.

Loan Maturity(Pre-Maturity) Indicator - Loan maturity refers to the final payment date of a loan at which point the principal is due to be paid off. As a loan approaches its maturity date, its remaining balance is small and the borrower more likely to pay off. The RM PD model uses pre-maturity indicator at 3 months, 6 months and 12 months to capture the loan's propensity to prepay when a loan is 3m, 6m and 12m away from being fully paid off. Please note the parameter estimate for 6m reflect an additive effect of both 3m and 6m parameter coefficients. Similar intuition follows for the 12-month indicator.

State Level Foreclosure Time- Depending on the state and circumstances, a foreclosure will be either non judicial or judicial. States which are in the judicial district have a prolonged foreclosure time. Any loan originated in a state with extended foreclosure timeline, tend to exhibit lower tendency to prepaid via short sale when a loan is in deep delinquency bucket.

Prepayment Seasonality - Seasonality in loan performance is a well-established characteristic of mortgages. Seasonality refers to any type of trend that is periodic in nature. For example, borrowers may be more likely to prepay in certain months than in others. House selling or buying, which results in housing turnover, usually is more popular in summer time while slows down in winter. The month specific seasonality has been captured in the model using monthly dummies (May, June, July and December).

S&P500 index - The S&P500 index is an American stock market index based on the market capitalization S&P500 large companies whose common stock are listed on the NASDAQ or NYSE. The S&P500 is one of the most commonly followed indices and is often considered the bellwether of the US equity market and the economy in general. The S&P500 index has been incorporated in the 2019 RM model at the recommendation of the model reviewers' to incorporate an attribute, which can track the effect of the domestic equity market performance on the NA mortgage portfolio. In order to capture the effect of the idiosyncratic risks of this index on the real estate performance, a magnitude variable was created which tracked the change in index value versus 3-month ago. As expected, an increasing index value is reflective of a stable economy which cascades down to people being optimistic about the market movements which lead to an increase in prepayment/turnover activities. For additional details, please review the PD model specification attachment referred to at the beginning and end of this response. Please note the inclusion of the S&P500 attribute has been made at the behest of the external regulators who recommended tracking the influence of equity market movements on the loss projections. Correlation analysis between this newly added SP500 attribute and final variables has been shared in Section 6.1 of the MDD, which provides the VIF measure, which in turn justifies the reasonableness of using such variable in the model.

HPI Appreciation (Logarithm) and its splines– The RM PD model leverages CoreLogic's Home Price Index, which shows home prices over the entire model development period. The VP transitions use the log of HPI appreciation (and its splines) since origination to factor in the influence of HPI for the VP movements. The log is an appropriate transformation to use since it can capture non-linear trends in the data spanning across different decades. Added justification around the HPI attribute and its importance can be found in the PD model specification attachment.

HPI interaction with MTM CLTV – The HPI attribute is interacted with marked-to-market CLTV to specifically capture the adverse impact on the borrower's equity during periods of decreasing HPI.

Higher values of MTM CLTV imply lower equity in the property, which decreases the propensity to prepay (ability to pay). This when coupled with periods of decreasing HPI now also affects the borrower's willingness to pay or upgrade their home as their overall confidence in the economy decreases.

B) Involuntary Prepay (IVP) transitions

Involuntary prepay or IVP reflects the loan propensity to default. The following attributes came up as significant predictors of the residential mortgage IVP movements across all transitions.

Loan-in-Trial Indicator- All loans that are in the trial period are less likely to IVP/default as they are actively locked down in the trial modification process.

Interest Only Loan- An interest only loan has a low monthly payment first few years after the initiation of the loan, since the borrower only pays the interest on the mortgage during the interest only period. When the interest only period ends, the scheduled payment significantly increases as the borrower will start to pay off the principal balance. Besides, some IO loans are also ARMs, so the rate goes up periodically based on the movements of the market rates. When the payment shock happens, the interest only loans are usually considered to be risker. And when housing prices fall, the IO loans are more likely to experience the underwater risk since the borrower don't build equity on the home via making principal payment. So borrowers with IO loans tend to default more than others and are positively correlated with IVP.

ARM Indicator – ARM loans, similar to IO loans are characterized by adjustable rates wherein the rates adjust to the changing market rates (Libor, Prime, etc.), once the initial low rate period is up. Hence all ARM loans are positively correlated with IVP

Lien indicator– Lien positions reflect the priority of claim on given collateral. First liens are considered less risky compared to all succeeding liens due to their *a* priori claim. Hence first liens have a lower probability to IVP compared to second or other junior liens.

GOV- Government loans are generally fall under the purview of the FHA (Federal Housing Administration) and VA (Veteran Affairs) programs, under which the FHA/VA approved lenders, are mandated to issue loans to borrowers, given that they meet certain credit criteria. Since the credit criteria and loan qualification requirements associated with the government loans are more relaxed compared to conventional loans; there is a mandated insurance requirement associate with these loans to protect the lenders from the increased credit risk. The presence of the compulsory insurance requirement reduces the probabilities of these moving to default/IVP.

PMI Indicator- Private Mortgage Insurance (PMI) is a special type of insurance policy, provided by private insurers, to protect a lender against loss if a borrower defaults on his/her loan. Most lenders require PMI when a homebuyer makes a down payment of less than 20% of the home's purchase price – or if the mortgage's loan to value (LTV) ratio is in excess of 80% (the higher the LTV ratio, the higher the risk profile of the mortgage). Borrowers typically tend to pay their PMI monthly until they have accumulated enough equity in the home that the lender no longer considers them high-risk. As expected, real estate loans with PMI coverage have a lower propensity to go IVP and hence have a negative correlation with the afore-mentioned dependent attribute as the credit losses are insured.

Channel Indicator(Wall Street) - The mode of origination – via a channel is reflective of a loan's credit quality based on the underwriting standards. Typically retail originations are presumed to have the highest loan quality over all other channels as these loans are originated by Bank's retail branches, adhering to stringent underwriting standards. The Wall Street channel refers to the bulk loan purchases made by Bank, prior to the crisis which included loans backed by the Wall Street investors. Most of these loans were sub-prime originations and hence had a higher propensity to go IVP/default across all liens. The Wall Street channel is further interacted with the second liens as second liens have a much higher propensity to charge off when in BUK6 compared to first liens.

Balloon Loan – Balloon loans are type of loans that do not fully amortize over their term. Since they are not fully amortized, a balloon payment is required at the end of the term to repay the remaining principal balance of the loan. Given the big lump sum payment due at the maturity, these loans tend to be charged off more likely compared to non-balloon loans and hence are positively correlated with IVP.

Foreclosure Time – The length of the foreclosure timeline affects the loan's propensity to ultimately go IVP/ default. Prolonged foreclosure times in certain states can lead to alternative workout arrangements like a short sale, modification or deed in lieu of foreclosure; all of which lowers the probability of the loan to ultimately default. The foreclosure was specifically introduced for the conventional loans, which reflects an additive effect. Additionally, a 12-month foreclosure time ratio is calculated specifically for the conventional loans, as the foreclosure process tends to differ for the government loans, which have specific loss mitigation strategies and as such, the foreclosure timeline has no effect on the government loans foreclosure processes.

FRM Loan (15 year or less): FRM loans with15 years or less term tend to have lower interest rates and demonstrate stronger credit characteristics of a given borrower (as they have strict qualification criteria in terms of income and debt ratios). With a lower interest rate, the payment of 15-year loan is more attributable to principal balance than a 30-year loan. Besides, borrowers with short-term loan are able to build up home equity more quickly than the longer terms. As a result, short-term FRM loans have a lower probability to default/IVP.

Principal Balance and its splines – A loan with larger unpaid principal balance usually take longer time to be charged off. Principal balances less than 5K have an increased propensity to be charged off rather than retaining on book as it is more cost-beneficial for Bank to charge-off these smaller amounts, instead of continuing to incur operational expenses on them. Principal Balance is interacted with the CMI vs (conventional vs government loans and lien position – first vs second liens), CPB entities to capture the idiosyncratic borrower profile between these entities. Typically, CPB portfolio caters to an affluent with higher net worth and corresponding higher loan balances. Borrowers with higher balances usually have a prolonged charge off timeline, due to policy dictates on the amount that can be charged-off.

Judicial State – Judicial states usually face a court adjudicated loan foreclosure process, which often leads the lenders to adopt alternate strategies like loan modification, short sale, etc. to save costs and expedite processes. Hence, judicial states have lower propensities to default or go IVP. The judicial indicator is interacted with the Conventional loans as Government loans follow a different IVP/loss mitigation policy once the loan is deeply delinquent and as such the judicial vs non-judicial split is not as relevant for the government loans.

Foreclosure Moratorium- As per Bank's foreclosure moratorium, Bank tries to not initiate or fulfill a foreclosure sale on any eligible borrower where Bank owns the mortgage, the borrower is seeking to stay in the home, which is his or her primary residence, is working in good faith with Bank and has sufficient income for affordable mortgage payments. Given this policy mandate, loans which fall under the purview of the foreclosure moratorium, have lower propensity to default. Please note, the foreclosure moratorium policy was put in effect prior to the establishment of the loan Modification program and hence only a few loans fall under the purview of the Foreclosure Moratorium policy.

Please note that during the observation window of model development data, there were two industry-wide events: 2009 mortgage foreclosure moratorium and 2012 national mortgage foreclosure settlement(http://www.ncsl.org/research/financial-services-and-commerce/national-mortgage-settlement-summary.aspx) which greatly impacted the mortgage foreclosure practices and the associated foreclosure timeline affecting the model performances, especially for loans in BUK7/IVP. The details on Bank's mortgage foreclosure moratorium process can be found in

https://www.Bankgroup.com/Bank/news/2009/090213a.htm

and

https://www.Bankmortgage.com/Mortgage/misc/Bank-Foreclosure_Suspension_ReleasePDF.pdf.
Calendar date related control variables (Foreclosure Moratorium and Foreclosure timeline) are created and tested to account for the impact of these two events on subsequent model performances. It is worth noting that 2012 national mortgage foreclosure settlement policy indicator was carefully tested but not included in the final charge-off model as it happened in the out-of-time period (01/2012 to 03/2014).

State Specific Dummy - State specific dummies (New York, Arizona, Florida, and New Jersey) have been inserted within the IVP equations to account for the state specific IVP/foreclosure related policy effects, which vary widely between states thereby affecting the loan's terminal status. Please note that state specific dummies are only interacted with conventional loans as government loans have specific loss mitigation strategies that are invariant on the state and its judicial system.

S&P500 Change for conventional loans - A improving economy (demonstrated by increasing index) leads to increasing market rates (Fed Funds, Prime rate, etc.). An increase in market rates lowers the values of fixed income securities leading to their sale in a high rate environment. This leads to an influx of cash which can be used to make missed mortgage payments thus lowering the chances to IVP. Please note the inclusion of the S&P500 attribute has been made at the behest of the external regulators who recommended tracking the influence of equity market movements on the loss projections. Correlation analysis between this newly added SP500 attribute and final variables has been shared in Section 6.1 of the MDD, which provides the VIF measure, which in turn justifies the reasonableness of using such variable in the model. The S&P500 attribute is transformed and is used as a 3-month ratio (current vs 3 months ago) for the conventional loans in the BKT7->IVP transition. While the model did test the sensitivity for other specific segments such as government, CMI vs CPB, the conventional loan segment remained relevant and significant in the course of model fine-tuning. All other interaction attributes were eventually dropped due to either unintuitive interpretation of insignificant estimate.

Unemployment rate for non-first liens – Rising periods of unemployment is associated with decreased levels of savings and correspondingly higher chances of missed payments leading to an increase in default/IVP. The unemployment rate is interacted with non-first lien loans which captures two-fold effect- first when unemployment rate increases, the borrower would make all efforts to make payments on his/her primary residence (first lien) thereby relegating non-first liens to a lower priority; second when unemployment rate is high, it lowers the possibility that an already delinquent borrower can recover due to lack of ability to pay.

HPI Appreciation since origination – Increasing HPI leads to more equity in the home due to rising house prices. Increased equity lowers the LTV, which increases borrower's confidence in economy and positively affects his/her willingness to pay, especially for borrowers who are already in deep delinquent buckets.

HPI change (12 months) for second liens – When home prices are declining over the 12 month period, generally as a policy measure, Bank will be more likely to charge off 150 DPD (bucket 6) second lien loans rather than let them continue rolling into deeper delinquency buckets since recoveries on second lien loans are anticipated to be lower relative to first lien loans.

CLTV and its splines - Higher CLTV implies relatively higher debt burden compared to the current property value. Hence higher CLTVs are always associated with higher default. Several CLTV attributes (including their interaction and spline) were included in the IVP related transitions as mentioned below.. Please note that the CLTV attribute only appears in the BUK6->IVP and BUK7-> IVP models, yet excluded in the BUK4->IVP, BUK5->IVP, and BUK7->IVP models. The reason why this attribute does not appear in the other IVP equations is primarily because they appear insignificant in these models and were dropped as part of the iterative variable selection process. For added justification on the spline for marked-to-market CLTV (spline flooring at 40 and capping at 120) for the BUK7->IVP, please refer to attachment -5.1.4 Bubble Plot Justification.xlsx. The MTM CLTV is also interacted with government loans and second liens. Government loans that are already in deep delinquency, are dictated more by the loss mitigation policies and less by the changes in borrower's CLTV. Second liens are usually charged off when in BUK6 than retaining them and letting them roll to BUK7 as recoveries made on second lien loans are relatively small compared to first liens.

Government_new_fcl – Tracks the liquidation policy change since Jan 2016. The business has faced challenges in government loan liquidation, as the process to liquidate government loans is not the same as conventional loans. Government loans are insured, and in order to file the insurance claim and obtain the insurance proceeds the property must be conveyed to HUD. As part of this process, Bank must meet HUD's guidelines for conveyance. These guidelines require eviction to be complete, as well as for the property to be in "conveyance condition." This is a higher standard than what is required for conventional loans, where there are no similar requirements around eviction and property condition. For conventional loans elements such as occupancy and property condition impact REO price but are not preconditions for a loan to be sold through REO. These problems in conveying government loans are caused by two factors: 1) changing and inconsistent interpretations of HUD requirements, impacting all servicers, and 2) idiosyncratic challenges Bank has faced in the management of HUD conveyance where the operational group has been transitioned between sites and has undergone multiple leadership changes. Very recently (starting around 2016) liquidation performance has improved due to increased management and regulatory scrutiny in this area. As the servicing of these loans transitions to Cenlar as part of the mortgage transformation sub-servicing initiative it is expected that these loans will continue to be serviced at least as well as they have been recently, as Cenlar is a top-rated HUD servicer.

C) Migrate To Worse (BUK01 -> BUK2 & BUK2 -> BUK3) transitions

Migrate to worse transitions are the intermediate and natural outcome of the transition framework which models the propensity of a loan to go worse the next month given its characteristics and performance status as of the source month. The following attributes came up as significant predictors for the residential mortgage PD migrate to worse movements across all transitions.

Channel Indicator (Wall Street, Correspondent, Broker) - The mode of origination – via a channel is reflective of a loan's credit quality based on the underwriting standards. The Wall Street channel refers to the bulk loan purchases made by Bank, prior to the crisis which included loans backed by the Wall Street investors. Most of these loans were sub-prime originations and hence had a higher propensity to go worse. Additionally, compared to the correspondent channel and the broker channel, loans originated via retail channel are expected to be less risky because of higher underwriting standards when the loans were originated.

Loan Purpose – An indicator denoting if the mortgage loan is either a refinance mortgage or a purchase money mortgage or a workout mortgage. Purpose may be the purchase of a new property or refinance of an existing lien (with cash out or with no cash out) or workout (loss mitigation strategy for borrowers in financial hardship). Compared to the other two types of mortgage, refinance purpose is the least risky because those borrowers have already demonstrated stronger performance in a higher interest rate environment and after refinancing, the mortgage are more likely to performing healthily. On the other hand, workout mortgage are the most risky ones, as it's for borrowers who historically cannot meet debt obligation and seek for relief via extending of loan term or rescheduling of repayment. 'Missing' loan purpose types are separately tracked to gauge their performance differential over time.

Loan Occupancy – the occupancy type is used to identify the customer intent for use of the property. This can be owner occupied, second home or investment. Typically properties which are not owner occupied (investment or second homes) are considered riskier as during any financial hardship, the natural inclination of the borrower would be to save the primary residence. Hence investment type occupancies tend to exhibit a higher probability to roll to worse.

Documentation Indicator- Higher/Full verification and documentation of borrower's income and assets are reflective of stringent underwriting standards which automatically weeds out the lower quality loans, at the time of origination. Hence loans with "full" documentation is less likely to go to delinquent. On the other hand, a loan with 'low' documentation such as borrower "stated income" or "stated assets" reflects lower underwriting quality with lower credit quality of the borrower and hence the elevated probability to roll to worse. 'Missing' documentation is separately tracked to gauge their performance differential over time.

PMI Indicator- Private Mortgage Insurance (PMI) is a special type of insurance policy, provided by private insurers, to protect a lender against loss if a borrower defaults on his/her loan. PMI is usually required on the conventional loan with less than 20% down payment. PMI is a layer of protection for lenders, but an added expense for borrowers. So borrower with the PMI burden are having higher propensity to roll to worse.

Co-borrower Indicator – Any additional number of borrower(s) adds a new layer of collateral protection which reduces propensity to go worse.

ARM Indicator, ARM reset in the past 6 month indicator, Post ARM indicator and Note Rate (ARM): – ARM loans are typically started out with a substantially lower interest rate than the more common fixed-rate loan. They have initial interest rates (or teaser rates) that stay low for a few years, depending on the contractor term. ARMs, which have a longer period of low rates, are less likely to go to worse. But when the initial low rate period is over the rates will adjust to the changing market rates (Libor, Prime, etc.) yielding a possibility of higher monthly interest payment and an higher risk to become delinquent, especially when reset just kicks in. But the reset risk is mitigated if an ARM can survive from the initial ARM reset period, after which the loan is more likely to continue to perform.

In addition, for ARM loans, it is expected the higher the interest rate, the higher the delinquency risk will be.

Interest Rate Spread at Origination (FRM) – The interest spread at origination reflects the yield spread premium on mortgage loans. The yield spread premium is often used to pay for the up-front loan origination costs, which are generally not passed on, to the borrower at the time of loan origination. As expected the higher the interest rate spread at origination, the higher the interest rate for the borrower and lower creditworthiness of the customer, which increase the borrower's propensity to roll to worse.

Current Note Rate(ARM) – In periods of rising rates, the ARM loans experience rising note rates which are generally indexed to some form of market rate. This makes the ARM loans sensitive to interest rate changes and corresponding payment shocks compared to FRM loans which have a fixed rate and are not impacted by rising rates

Interest Only Loan, Post IO indicator and P&I Payment Ratio (IO): In the first few years after origination, borrowers with IO loan only need to make the interest payment hence their payment burden is lower than the non-IO loans. So they are less likely to migrate into deeper delinquency buckets. However, when IO period ends, borrowers also need to make principal payment every month. Such a significantly increased payment will increase the delinquent risk. Payment shock variables by comparing the monthly payment at the observation month and 6 months ago was created to measure the payment shock associated with IO loans. The stronger the shock is, the more likely the loan will go to delinquent. Several payment shock variables are used in the model which correspond to – the following. Please review the specifications file for additional details.

B_S_INT_ONLY_IND	Indicator(IO loan)
POST_IO_IND	Indicator(entered IO reset period)
P_AND_I_ratio_IO	Payment (UPB and interest) ratio of current vs 6 months ago for IO
POST_ARM_IND	Indicator(entered ARM reset period)

Principal Balance and interaction with CMI loans - Higher principal balance is an indicator of higher mortgage payment obligation, or lower portion of the mortgage that has been paid down. So it become more difficult for borrowers to make the mortgage payment when they are experiencing financial hardship. On the other hand, borrowers who have paid-down most of their mortgage and with very low remaining balance (suck as a 10K or even less) may think that they have paid off their loans and stopped making future payment. The phenomenon is also visible is figure 4.1.2, where loans with almost zero CLTV are more likely to go delinquent. The principal balance is interacted with CMI loans to capture CMI (conventional only as government loans have specific loss mitigation strategies when a loan goes delinquent) specific balance effect.

Balloon Loan - Balloon loans are type of loans that do not fully amortize over their term and hence a balloon payment is required at the end of the term to repay the remaining principal balance of the loan. Balloon loans are generally considered to be of higher risk. Given the lump sum payment due at the maturity, these loans tend to roll to worse more compared to non-balloon loans.

Balloon Mature - Balloon mature indicator indicates that the loan’s MOB is greater than the loan term. Since the loan still exists on the book, beyond the stipulated loan term, reflects the credit riskiness of the loan and its inability to make the balloon payment. Hence these loans have a heightened probability to roll to worse.

Second lien – When a borrower faces financial distress, he/she is predisposed to using all means to make payments on the primary residence, which is also the first lien on the property. By virtue of seniority, second liens are more probable to roll over to worse, given the financial constraints of the borrower.

FICO – FICO score is an important characteristic that reflects the financial status of the borrower at a given point in time. Refresh FICO refers to the FICO that is reflective of the borrower's current credit conditions. Intuitively, higher the FICO, the better positioned the borrower is to fulfil his credit obligations on a revolving basis (both secured and unsecured). A lowering of FICO score is associated with deteriorating credit situation of the borrower which can make him/her to miss payments and roll to a worse DLQ bucket. The Refresh FICO and its splines has been used in the BUK01->BUK2 transition. For justification on the splines, please refer to the attachment -5.1.4 Bubble Plot Justification.xlsx. All splines of Refresh FICO have to be interpreted based on the additive effect as described earlier in the document.

Katrina/Sandy Impact- Borrowers affected by natural disasters are more likely to roll to worse given the extensive damages they incur on their existing property, health and personal belongings. Delay from subsequent insurance claims and subsequent recoveries usually follow a long course of action which further impedes the borrower to remain current on their monthly mortgage payments. Based on model sponsor leadership requests, effects from natural disasters such as hurricanes were included in the list of risk drivers considered in the 2019 model redevelopment project in the form of indicator variables.

Age of oldest trade line- Borrowers with a long credit history are generally considered as high-quality credit-worthy customers. The delinquency risk is expected to be lower when credit history is longer.

Post Crisis Origination- Loans originated post crisis (post 2010) reflect stringent policy requirements with improved underwriting standards. As such, these loans are deemed as likely to roll to worse. On the other hand, government loans that originated post crisis, demonstrate higher likelihood to roll to worse as these loans

Ever 60DPD Indicator – This variable reflects whether a loan ever experienced 60DPD and is used in BUK2 model. Some of the borrowers that rolls into BUK2 are the ones who forget to make monthly payment. They are more likely catch up with the schedule mortgage payment and stop rolling into deeper delinquent bucket. The ever 60DPD indicator is used to separate those sloppy payers from the ones who had a real dirty history and are more likely to roll up.

CLTV and its splines: As explained earlier, higher CLTV implies relatively higher debt burden, combining all liens, as encumbered on the given property. Hence higher CLTVs are always associated with higher propensity to roll to worse. CLTV effects are observed to be different across lien position (first versus second), and across loan type (conventional and government). The rationales for interacting the CLTV attribute with Government and lien position has been described before in the –to-worse transitions justifications. For added justification on the spline usages, refer to attachment - 5.1.4 Bubble Plot Justification.xlsx

Unemployment Rate - Increased unemployment increases the propensity of a loan to roll into worse due to the increased financial hardship on the borrower. Unemployment rate is deployed depicting the 12-month change, specifically for the CMI, IO Conventional, Government and Broker vs Correspondent channel. The 12-month change captures the decreased ability to pay due to increase in unemployment in the short term. The decreased ability to pay stems from homeowners unable to make payments due to lack of income and their difficulty to relocate. The effect is particularly strong for the IO borrower because IO borrowers, who choose to make interest only payment in the first few years, are usually the ones who are more payment sensitive or income constrained when unemployment rates start to rise. Government loans have a delayed response to unemployment rate because government loan borrowers in general have lower debt to income ratio compared to conventional loan borrowers after controlling for CLTV and FICO. Govt loans have more provisions on income source, are stricter about debt-to-income ratio and what sources to pay for the loan. On the other hand, conventional loans require bigger down payments and lower monthly payment whereas government loans require smaller down-payments and have higher monthly payment. Borrower who took out GOV loans is likely to be more optimistic about their future income due to self-selection. The unemployment rate is also segregated by channel - broker and correspondent, as these channels demonstrate differentiated sensitivities to the 12 month unemployment rate change.

HPI (Splines and interaction w Wall Street) - Increasing HPI leads to more equity in the home due to rising house prices. Increased equity lowers the LTV which leads to lower propensity to roll to worse. The sensitivity can be stronger in the Wall Street channel, which are composed by more pre-crisis sub-prime loans. The 12 month HPI change is interacted with CMI and CMI - Conventional to capture borrower's lack of willingness to pay due to lack of confidence in housing market, home price appreciation and the economy. Further, the HPI appreciation is tracked for the CMI portfolio, which is capped at 1, which reflects the idiosyncratic risk-taking behavior of borrowers who speculate on short term home price appreciation and are more likely to stop payments when house prices are no longer rising or no gains from purchase. An indicator is also created for conventional loans which have already survived a period of decreasing HPI. Borrowers that have survived from period when home value has dropped below purchase price have demonstrated their resilience to home price depreciation and negative macroeconomic shocks. These borrowers have lower propensity to roll to worse.

MOB and its interaction with Wall Street Channel – As mentioned earlier, Wall Street originations are typically considered riskier, reflecting the pre-crisis sub-prime originations. However if such a loan is considerably seasoned then they reflect a borrower's good credit standing and are less likely to roll to worse. Among the seasoned loans, the initial risk that is associated with origination channel has been diminished significantly. As such, to not over penalize these WALL St customers, a time-decay factor denoted by Month on Book has been incorporated for the Wall St loans.

HPI Burn-in Effect – This measures the loan's tenacity to survive a period of declining HPI which demonstrates its resilience to macroeconomic shocks/business cycles, given its strong credit profile. Such a loan is assumed to have lower propensity to roll to worse, other things being kept constant.

Month Specific Dummy - Seasonality in loan performance is a well-established characteristic of mortgages. Seasonality refers to any type of trend that is periodic in nature. For example, borrowers may be more likely to miss payments in certain months than in others. In February or March, a loan is less likely to go to delinquent because borrowers get extra cash from annual bonus or tax return to make their mortgage payment.

Income 12-month ratio – Rise in income implies increased spending power, which in turn lowers the propensity to go worse as it positively impacts the borrower's ability to pay.

D) Migrate To Better (Fully Cure & Partially Cure) transitions

Migrate to Better transitions are the intermediate and natural outcome of the transition framework which models the propensity of a loan to either fully cure or partially cure the next month given its characteristics and performance status as of the source month. The following attributes came up as significant predictors for the residential mortgage PD migrate to better movements across all transitions.

Fully Cure (Roll Back to BUK01) transitions

Co-borrower Indicator – Any additional number of borrower(s) adds a new layer of collateral protection which increases the propensity of a loan to fully cure.

Loan-in-Trial Indicator- All loans that are in the trial period are less likely to cure as the trial process is itself induced based on the borrowers' financial hardship which led to a series of missed payments transitioning the borrower to a worse DLQ bucket each time

ARM Indicator, ARM reset in the past 6 month indicator, and Note Rate (ARM):– ARM loans are characterized by adjustable rates wherein the rates adjust to the changing market rates (Libor, Prime, etc.), once the initial low rate period is up. Before an ARM loan resets, it is more likely to cure as it is associated with a lower rate and thus lower monthly payment. However, once the ARM loan gets reset, it would experience a payment increase which makes it riskier, especially within recent past reset period (e.g. within post-reset 6 months). And for ARMs, the dynamically changing interest rate is negatively associated with the curing rate.

Channel Indicator (Wall Street, Correspondent and Broker) - The mode of origination – via a channel is reflective of a loan's credit quality based on the underwriting standards. The Wall Street channel refers to the bulk loan purchases made by Bank, prior to the crisis which included loans backed by the Wall Street investors. Most of these loans were sub-prime originations and hence had a higher propensity to go worse. Additionally, compared to the correspondent channel and the broker channel, loans originated via retail channel are expected to be less risky because of higher underwriting standards when the loans were originated.

Interest Only Loan- Borrower who takes an interest only loan only need make small amount of interest payment in the first few years, after which he will start to make significantly increased principal payment every month. In either period of the IO life time, before or after reset, a deep delinquent represents the borrower's financial hardship in making the mortgage payment. Compared to non-IO borrowers, the possibility for them to be cured is expected to be lower.

Lien indicator – Lien positions reflect the priority of claim on given collateral. Second liens are less likely to fully cure as the borrower is more likely to prioritize his first lien (to prevent loss of home ownership) during situations of financial hardship.

Loan Purpose – As expected, workout mortgages reflect higher risk compared to other mortgages and hence are less likely to completely cure.

Property Type- Compared to single family home (SF), condo and multi-family property is the preferred choice of borrowers with limited financial ability. Those borrowers are less attached to the property than the owners of the SF. In addition, condo price volatility is higher than SF. Once delinquent, those borrowers are less likely to cure.

CPB – CPB portfolio, given its affluent client mix is expected to fully cure.

Principal balance and its splines – Principal balance reflects payment burden. Therefore, once being delinquent, higher principal balance, which indicates higher payment burden of a loan, less likely to fully cure.

Foreclosure time – For loans in BUK7, the longer the foreclosure time, the lower the chances that the loan will fully cure. This is because usually the longer a loan stays in BUK7, the less likely it can be cured. The extended foreclosure time prolongs the portfolio average month in BUK7, hence are negatively associated with BUK7 curing rate.

Government Loan – Government loans have relaxed underwriting criteria and hence are less susceptible to fully cure.

HPI and its splines & interaction with MOB effect– Increasing home prices, lowers borrower loan amount compared to accumulated equity in the property, leading to higher chances of fully curing. But such a HPI sensitivity are much lower for loans that have been paying for a long time say, more than10 years. Deep delinquency loans with loan MOB usually means that the borrowers are experiencing financial hardship. Even when HPI improves, they may still don't have the ability to pay the overdue payment.

Income Ratio – Increasing state level income are reflective of a booming economy leading to higher chances to cure.

CLTV and its splines - Higher level of CLTV leads to increasing cumulative debt burden (considering all liens on a given property) leading to lower chances of fully curing.

MOB on new loans (MOB<=36) – Borrower who just started a mortgage may have a higher propensity of forgetting to make the monthly payment and then rolls into delinquent (30~59+DPD). Those "sloppy payer" are not the real financially troubled customer so they are more likely to cure. In this case the curing possibility are negatively associated with the age of a loan. But such an effect only happens on the relatively new loans. Among borrowers with long enough mortgage experiences, curing rate no longer related to loan age.

Unemployment rate and its interaction with period of falling HPI – Increasing levels of unemployment leads to lower chances of loan fully curing due to financial hardship on the borrower. In fact, if unemployment rate is accompanied by declining housing prices, it further elevates the debt burden on the borrower leading to further lowering the propensity to fully cure

Debt Ratio – The debt ratio reflects the borrower's overall revolving debt(secured + unsecured) relative to his/her household income and higher the ratio, less chances to fully cure

Interest Rate Spread at Origination – Higher original spreads indicate that the loan was issue at above market rate to given the risky profile of the borrower. As such, loans with note rate above par indicate that the borrower is deemed to be risky and hence had fewer chances to cure back from delinquent status.

Seasonality Effect - Seasonality in loan performance is a well-established characteristic of mortgages. Seasonality refers to any type of trend that is periodic in nature. For example, borrowers may be more likely to miss payments in certain months than in others. Usually holiday season would result in missing payment or less likely to fully cure. Months past holiday season, like February or March, a loan is more likely to go back to current bucket.

Ever 60DPD Indicator –. Some of the borrowers that rolls into BUK2 are the ones who forget to make monthly payment. Compare to those who have a dirty history, they are more likely catch up with the schedule mortgage payment and roll back to current bucket.

Partially cure transitions

Co-borrower Indicator – Any additional number of borrower(s) adds a new layer of collateral protection which increases the propensity of a loan to partially cure.

Trial Indicator - All loans that are in the trial period are less likely to partially cure as the trial process is itself induced based on the borrowers' financial hardship which led to a series of missed payments transitioning the borrower to a worse DLQ bucket each time.

ARM Indicator, ARM reset in the past 6-month indicator, and P&I ratio (ARM)– ARM loans are characterized by adjustable rates wherein the rates adjust to the changing market rates (Libor, Prime, etc.), once the initial low rate period is up. Before an ARM loan resets, it is more likely to cure as it is associated with a lower rate and thus lower monthly payment. However, once the ARM loan gets reset, it would experience a payment increase which makes it riskier, especially within recent past reset period (e.g. within post-reset 6 months). In addition, the greater the payment shock for an ARM, the lower the probability that the loan will partially cure.

Documentation – Loans with FULL documentation type demonstrate stronger credit profile of the borrower and are more likely to partially cure compared to loans with LOW documentation type or MISSING doc type which mainly applies to very old loans or acquisitions.

Interest Only Loans – In either period of the IO life time, before or after reset, a deep delinquent IO indicates the borrower's financial hardship in making the mortgage payment. Compared to non-IO borrowers, the possibility for them to be cured is expected to be lower.

Lien Indicator – Second liens are less likely to partially cure as the borrower is more likely to prioritize his first lien (to prevent loss of home ownership) during situations of financial hardship.

Occupancy Type: Secondary homes/Investment properties are less likely to partially cure than owner occupied homes as the primary residence usually get a priority of payments during financial constraints /hardships.

Wall Street Loans – Loans originated through the Wall Street channel have elevated riskiness and are less susceptible to partially cure.

Indicator for Current HPI being lower than Origination HPI – Properties that are currently experiencing a period of declining HPI are less susceptible to partially cure.

Principal Balance and its splines – Principal balance reflects payment burden. Therefore, once being delinquent, higher principal balance less likely to partially cure. Such an effect is more prominent when under stressed macroeconomic environment.

FRM(15 yr. and less) – Fixed rate loans with term less than 15 years demonstrate stronger credit characteristics of the borrower hence they are more likely to partially cure, if they become delinquent.

Government Loans – Government loans are usually originated with lower credit quality and thus less likely to partially cure.

HPI Appreciation and its splines – Growing housing prices are reflective of increased home equity leading to increased chances of partial curing. Similarly, periods of falling HPI (growth rate <1) lowers the chances of partial curing.

Year over year Income change– Borrowers are more prone to make extra payments and partial cure when the overall state level aggregated income increases.

S&P500 – Improving equity markets are reflective of a booming economy which translates to increasing housing prices (indirect effects) and more equity in the home leading to the borrower making payments on a delinquent loan. Please note the inclusion of the S&P500 attribute has been made at the behest of the external regulators who recommended tracking the influence of equity market movements on the loss projections. Correlation analysis between this newly added SP500 attribute and final variables has been shared in Section 6.1 of the MDD which provides the VIF measure, which in turn justifies the reasonableness of using such variable in the model.

CLTV and its splines – Higher level of CLTV leads to increasing cumulative debt burden (considering all liens on a given property) leading to lower chances of partial curing. Please note the CLTV spline represents an additive effect.

MOB and its splines – The longer a loan is on risk book the more likely it stays in the same delinquent bucket, compared to going to worse.

Origination Interest Spread – The higher the spread, the higher the note rate over the market rate at origination. This is reflective of a risky borrower (hence the added spread to compensate the lender) who are less likely to partially cure, if delinquent.

Unemployment Rate – Increasing levels of unemployment leads to lower chances of loan partially curing due to increasing financial hardship on the borrower.

Debt Ratio - The debt ratio reflects the borrower's overall revolving debt (secured + unsecured) relative to his/her household income and higher the ratio, less chances to partially cure.

Annualized MoM GDP change: A increasing GDP is indicative of a healthy economy in which a loan is more likely to partially cure.

Age of oldest mortgage- Borrowers with a long credit history are generally considered as high-quality credit-worthy customers. The possibility of curing is expected to be lower when age of oldest mortgage is longer.

Number of open mortgages- Curing rate are expected to decreases with number of mortgages, because borrowers with more mortgages are usually having heavier credit obligation and they are impacted more negatively when house price drops. It will be more difficult for them to be cured once delinquent.

The Residential Mortgage PD model specification can be found in the attachment 5.1.4 Residential Mortgage PD Model Specification.xlsx. For detailed parameter estimates, significance, or VIF, please refer to the attachment '6.1 Residential Mortgage Diagnostic and Statistical Tests.xlsx.'

LGD Model

The pool of candidate variables for the Loss Given Default (LGD) model varies between the stages modeled wherein the first stage models the losses based on their outcome [full vs partial vs zero losses] and the second stage models the loss severity for the partial losses only. As already mentioned in Sections 3.1 and 5.1.2, there has been a change in modeling framework for the first lien loans from prior methodology. The new approach aligns the loss calculation methodology across first and second liens. All rationale outlining the change in the modeling approach with supportive justification has been outlined in Chapter 3 of the MDD.

The LGD model similar to the PD model includes static loan and borrower attribute variables such as business entities, macro-economic variables such as home price growth rate, unemployment rate; dynamic variables such as updated LTV, balance, loan age, etc. While the candidate variable definitions' are same as explained in the PD listing, the intuitive explanation based on the parameters' magnitude and direction are illustrated as below, based on the stages modeled. The first stage models the binary logit regressions (Full vs Partial and Zero vs Partial) and the next stage models the probability of partial losses using a logistic regression.

First Stage Model (Full vs Partial and Zero vs Partial)

The first stage – first lien model is a binomial model which models the propensity of zero losses or full losses using partial losses as a reference.

Marked-to-Market Loan-to-Value (MTM LTV) - LTV measures the underlying equity in the property and hence it's a measure of the amount of financial leverage that pertains to a given property. For new loans, the original LTV reflects the borrower's down payment as a percent of property value, which in turn captures borrower's commitment to the property as well as borrower's long-term financial capacity. MTM LTV is calculated as the ratio of updated current loan balance divided by updated house price using the county level housing price change (as measured by the licensed CoreLogic index which reflects the loan's amortization. A detailed discussion on the CoreLogic update has been added in Chapter 5, Section 5.1.1 of the MDD). This provides a more accurate time-varying measure of borrower equity in the property. Higher MTM LTV leads to relatively lower zero losses compared to partial losses.

Current Balance and its Splines - This attribute is expected to be negatively correlated with severity, mostly likely reflecting the larger loan balances (jumbo loans), which by virtue of Bank's stringent underwriting standards are reflective of better risk profiles. Loans with relatively larger current balances tend to fall less under the realm of full losses using partial losses as reference.

Housing Price Index Growth Rate (12 months) - The Housing Price Index (HPI) is a weighted and repeat-sales index that can provide timely and accurate measure of house price trends at various geographic levels. The HPI change is a broad measure of the movement of house prices change over a given time period. Various home price change variables are included in the equations such as most recent 12 months home price change, peak to current and trough to current home price change in the last 3 years, etc. Positive HPI change suggests home value appreciation, and is associated with lower loss severity. A rising HPI leads to relatively more zero losses and less full losses using the partial losses as reference.

Unemployment Rate - It measures the prevalence of unemployment in the economy and is calculated as a percentage of the number of unemployed individuals by all individuals currently in the labor force. During periods of recession, an economy usually experiences a relatively high unemployment rate. Both unemployment rate and its rate change are important loss severity model equations. High unemployment rate and (or) unemployment leads to more full losses.

Judicial vs Non-judicial - Essentially, there are two types of foreclosure procedures; judicial foreclosure and non-judicial foreclosure. Generally, states that use judicial foreclosures use the court system to execute the foreclosure; states that use deeds of trust conduct non-judicial foreclosures, using an out-of-court procedure defined by state law.

Typically borrowers in judicial states tend to linger in the mid-to-high delinquent buckets longer than going into default due to the extended foreclosure timeline as part of the legislation process. Further to the specific judicial laws, the borrowers are less likely to consider other workout options and hence are less likely to partially/fully cure or even payoff completely. In the judicial states, borrowers generally sign two separate instruments: the note (or bond), which is evidence of the borrower's promise to pay the debt; and the mortgage, which is the legal instrument that creates the lien on the property as security for the debt. If the borrower cannot pay the mortgage, the lender hires an attorney, who begins legal action to protect the lender's interest. The attorney files several documents: a summons directing the borrower (defendant) to appear in court and formally begin the legislation process. The documents are then filed with the clerk of the court in the county where the property is located.

On the contrary, in the non-judicial states, foreclosures are based on deeds of trust that contain a power-of-sale clause. The clause enables the trustee to initiate a foreclosure sale of the collateral (home), without having to file a lawsuit or go to court. The trustee is typically required to issue a notice of default and notify the trustor (borrower) accordingly about the defaulted loan status. If the trustor does not respond, the trustee is entitled to initiate the steps for conducting the foreclosure sale of the collateral (home).

Loans that belong to a judicial jurisdiction tend to experience relatively more full losses. In the Method A first lien – first stage model, the judicial indicator is being interacted with the diminishing unemployment rate, to gauge the effect of diminishing unemployment trends on loss severity for a loan that falls under the purview of a judicial state. As expected, with falling unemployment, judicial states tend to riskier and experience more full losses compared to loans in non-judicial states.

Combined Loan to Value – The CLTV ratio measures the total equity that a borrower has in a secured borrowing based on current estimates of the value of the collateral and considering all lien positions related to the given property. In the second lien, stage 1 model, the CLTV attribute is interacted with the unemployment rate to gauge the effect of increasing CLTV on the loan losses in a falling unemployment environment. As expected, higher CLTVs lead to lower zero losses and more partial losses.

Residential Mortgage Indicator –Typically when a property goes into foreclosure, the lender holding the home equity loan does not get paid until the first mortgage lender is paid. Consequently, the home equity loan lender's risk is greater with lower zero losses compared to partial losses.

Junior Ratio – This is defined as the proportion of the second lien balances to the total balances whereas the total balances are defined as the sum of the first lien and second lien balances. Higher percentages of this ratio reflects an increased proportion of the second lien debt on the equity of the property which reflects a heightened riskiness when the property devalues due to falling housing price index or worsening economic trends. A higher value of the junior ratio leads to more full losses compared to partial losses as reference.

Lender Placed Insurance (LPI) - Mortgage loan agreements include a requirement that the borrower maintain insurance to protect the property serving as collateral for the loan and, if the borrower fails to maintain the required insurance or fails to provide required evidence of insurance, the lender, through the servicer, may place insurance on the property serving as collateral for the loan and charge the borrower for this insurance. LPI is a master insurance policy issued to the mortgage servicer as the policyholder and insured. The LPI insurance policy provides that coverage begins on any property in the servicer's covered mortgage loan portfolio at the instant that the borrower's voluntary policy ceases to provide the required coverage. The presence of LPI coverage indicate lower propensity for full losses.

Principal Reduction - A principal reduction is a decrease granted toward the principal owed on a loan. A principal reduction can be obtained to decrease the outstanding principal balance on a loan and provide relief for a borrower. Principal reduction is normally deployed to prevent foreclosures on properties, which may be more costly to financial institutions than a reduced principal owed to them. Borrowers who fall under the purview of the principal reduction program are typically characterized with some form of financial hardship and are generally correlated with full losses.

Second Stage Model

The second stage of the LGD framework models the partial losses in a regression modeling framework. Zero/Full losses are kept outside the purview of the second stage model.

Current Balance - Loss severities are also affected by loan size. For all vintages, the smallest loans exhibit the highest severity, and the severities decline monotonically with loan size. One possible reason is that loan size influences liquidation costs. A smaller loan is more likely to trade at a higher foreclosure discount due to the relatively larger fixed costs associated with exercising the foreclosure option. In addition, liquidation costs are likely larger on smaller loans as a percentage of loan unpaid principal balance (UPB).

Marked-to-Market LTV and its splines - MTM LTV is calculated as the ratio of updated current loan balance divided by updated home prices using the most up-to-date housing price index change. Higher values of the MTM LTV imply lower equity in the property and are usually associated with higher loss severities.

Private Mortgage Insurance (PMI) - Private Mortgage Insurance (PMI) is a special type of insurance policy, provided by private insurers, to protect a lender against loss if a borrower defaults on his/her loan. Most lenders require PMI when a homebuyer makes a down payment of less than 20% of the home's purchase price – or if the mortgage's loan to value (LTV) ratio is in excess of 80% (the higher the LTV ratio, the higher the risk profile of the mortgage). As expected, real estate loans with PMI coverage have a positive correlation with loss severity as these loans represent reflect financial hardship of the borrowers. As part of the modeling process, the insurance amount is added back to the cumulative losses to the loss data.

Judicial vs Non-judicial state – As mentioned before, judicial states tend to have prolonged foreclosure processes once a loan defaults which increases the loss severities associated with these states compared to non-judicial states.

Broker Channel - A mortgage broker acts as an intermediary who brokers mortgage loans on behalf of businesses. The broker performs some of the mortgage loan processing functions but it usually does not involve underwriting the mortgage loan, fund the mortgage loan at settlement, or service the mortgage loan. The mortgage loan is typically funded by the mortgage loan lender that commissioned the broker's services. Broker originated loans typically lack the stringent underwriting criteria associated with retail channel loans and are considered to be more risky with increased loss severities associated with them in the event of default.

Principal Reduction – A principal reduction is a decrease granted toward the principal owed on a loan. A principal reduction can be obtained to decrease the outstanding principal balance on a loan and provide relief for a borrower. Principal reduction is normally deployed to prevent foreclosures on properties, which may be more costly to financial institutions than a reduced principal owed to them. Borrowers who fall under the purview of the principal reduction program are typically characterized with some form of financial hardship and are generally correlated with higher loss severities in the event of default.

First Mortgage Balance – First mortgage typically represent the primary loan that pays for the property and has priority over all other liens and claims on the property in the event of default. Intuitively, in the event of default, the extent of loss severity incurred on second liens are dependent on the underlying unpaid balances due on the first liens. Since first liens represent the senior most liens, any subsequent recoveries made on the property tend to compensate the first liens in order of seniority. If the second lien tends to have a relatively large first lien balance, the loss severities would tend to be higher.

Mark-to-Market LTV – The marked-to-market LTV reflects the updated current balance due on the property as a proportion of the updated property value. The higher the MTM LTV, the increased debt burden of the borrower which leads to a higher loss severity in the event of default.

Unemployment Rate – Unemployment rate is generally considered an important determinant of mortgage performance and credit riskiness of a given loan. All things being kept equal, increasing levels of unemployment indicate borrower hardship leading to missed payments and increased default incidences and loss severities.

Mark-to-Market CLTV - The combined loan-to-value ratio is the ratio of all loans secured by a property to the property's value. Lenders use the CLTV ratio to determine the risk of default by prospective homebuyers when more than one loan is used. Marked-to-market CLTV reflects the updated balance due for all loans on a property divided by the updated proper value. Loans with higher overall MTM CLTV tend to exhibit higher loss severities due to the increased debt for the borrower compared to the equity in the property.

Deficiency Judgement – Defined as the judgment for the amount a homeowner owes the lender after a house is foreclosed upon and sold by the creditor for less than the actual that is secured by the asset. In most states, the lender can file a separate lawsuit to recover a deficiency owed by the borrower. Some states restrict deficiency judgments after a home foreclosure. For the states with deficiency judgement, due to the lawsuits filed to recoup the additional balances owed, the loss severities are deemed to be higher.

Residential Mortgage Indicator – Residential mortgage loans are traditional mortgages which are taken to purchase a property compared to a home equity loans which is taken out after there is some equity in the property. When the property goes into foreclosure, the lender holding the home equity loan does not get paid until the first mortgage lender is paid. Consequently, the home equity loan lender's risk is greater with relatively higher loss severities compared to a traditional residential mortgage.

For additional details on the LGD model specifications, refer to the attachment '5.1.4 Severity_model_specifications.xlsx.'

Are the variable transformations meaningful?

Yes. All Residential Mortgage model suite transformations were performed where appropriate in order to capture non-linearity and ensure model accuracy. Detailed data transformations used during model development have been described in Section 4.1.6 of the MDD with several important variable transformation examples were provided in earlier part of this section to justify their meaningfulness.

What were the variables originally considered? Which ones were provided by the business? Provide answers separately for risk and business drivers

All variables considered as part of the initial candidate pool were backed by economic intuition, business rationale, industry specific trends and loan level performance history.

How were the drivers (both risk and business) translated (or proxies) into model variables?

The drivers were tested as independent variables to the model and selected based on the criteria detailed in the above paragraphs.

Which variables were statistically significant and which were statistically insignificant? What is the evidence supporting this selection? Provide answers separately for risk and business drivers

For the 2019 Method A Residential Mortgage PD and LGD models, all finalized variables selected are statistically significant. Final models do NOT include any statistically insignificant variables. Because of the complicated structure of the model, the modeler did not explicitly list the variables that were insignificant and removed. As summarized in the attachments – '5.1.4 Residential Mortgage Model Specification' and '5.1.4 Severity_model_specifications.xlsx' and for each PD equation and LGD equation, modelers provided the approximate number of candidate variables that were removed / added / changed subject to economic intuition and statistically insignificance.

How do the drivers relate to the model outputs? Please explain. Provide answers separately for risk and business drivers

[Provide supporting evidence for the relationship between the chosen variables and the output based on qualitative and/or quantitative arguments. Include a list of alternative drivers considered that were not selected in the process with arguments for rejecting them.]

Please see the summarized table below, which links the key risk drivers including loan level, attributes, bureau specific variables(where applicable), macro-economic indicators to key outcomes for both LGD and PD models respectively. A detailed explanation linking the chosen variable to the output has been provided above. To summarize, all business/loan-level/macro-economic risk drivers identified below demonstrate sufficient add-on effect to the modeled equations. Attributes within the initial candidate pool had included an exhaustive list of all possible loan-level credit risk drivers. As part of subsequent model development, attributes, which turned out to be insignificant, or had counter-intuitive business intuition, were eventually dropped. Please be aware that variables were not excluded all at once, rather an iterative process was conducted to understand and validate the incremental add-on effect of each attribute to the equation. Please refer to attachments – '5.1.4 Residential Mortgage PD Model Specification.xlsx' and '5.1.4 Severity_model_specifications' for RM PD and LGD model specifications.

Table 5.1.4.5.1 PD Model Summarized Specifications

Key Risk Drivers	*To VP*	*To IVP*	*To Worse*	*To Better*
Loan Level Attributes				
FICO	**Positive**	**Negative**	**Negative**	
CLTV/LTV	**Negative**	**Positive**	**Positive**	**Negative**
Co-Borrower			**Negative**	**Positive**
In-Trial	**Negative**	**Negative**		**Negative**
Loan Documentation			**Miss>Low>Full**	**Full > Low >Miss**
Channel		**Wall St > Other**	**Wall St>Corr >Broker>Retail**	**Retail>Broker>Corr>Wall St**
GOVT Loan		**Negative**	**Positive**	**Negative**
Loan Purpose			**Workout > Purchase>Refi**	**Other>Missing>Workout**
Sandy/Katrina			**Positive**	
Prior DELQ Status				
Principal Balance	**Positive**	**Negative**	**Positive**	**Negative**
Judicial	**Negative**	**Negative**		
Lien	**first > 2nd**	**2nd+ > first**	**2nd+ > first**	**first > 2nd**
IO vs Non-IO		**IO > Non-IO**	**IO > Non-IO**	**Non-IO > IO**
FRM vs ARM	**ARM > FRM**	**ARM > FRM**	**FRM > ARM**	**ARM>FRM**
Occupancy Type			**Invest, Secdhome > Owner Occ**	
Vintage (post 2010 Originations)			**Negative**	
PMI		**Negative**	**Positive**	

Property Type	**Condo > Other**			**SFR>Condo > Multi**
Balloon		**Positive**	**Positive**	
Community Loan	**Negative**	**Negative**	**Positive**	
CPB Indicator			**Negative**	**Positive**
Foreclosure Moratorium		**Negative**		
Months in BUK7	**Negative**	**Positive**		
Origination Note Rate				
MOB			**Only Positive within 36M**	**Positive**
Reset(ARM/IO)	**Positive**	**Positive**	**Positive**	**Negative**
Current Note Rate			**Positive**	
Market Rate		**Positive**		
Interest Spread	**Positive**		**Positive**	**Negative**
Debt Ratio				**Negative**
Prepayment Penalty	**Negative**			
Foreclosure Time	**Negative**	**Negative**		**Negative**
Interest rate burnout effect	**Negative**			
Macroeconomic Attributes				
HPI	**Positive**	**Negative**	**Negative**	**Positive**
Unemployment Rate	**Negative**	**Positive**	**Positive**	**Negative**
Income				**Positive**
S&P500	**Positive**	**Negative**	**Negative**	
GDP				**Positive**
Bureau Attributes				
Age of oldest Mortgage				**Positive**
Number of open Mortgages				**Negative**
Age of oldest Tradeline			**Negative**	

Table 5.1.4.5.2 Severity Model Summarized Specifications

Attributes of Importance	*First Stage **		*Second Stage **
	Zero vs Partial	*Full vs Partial*	*Partial loss only*
Loan Level Attributes			
Current Balance (Inc. splines)		**Negative**	**Negative**
Judicial		**Positive**	**Positive**
HELOC/FRHEL Indicator	**Positive**	**Positive**	
Marked to market LTV	**Negative**		**Positive**
Marked to market CLTV			**Positive**
CLTV * Declining Unemployment	**Negative**		
Residential Mortgage Indicator	**Positive**	**Negative**	**Negative**

Attributes of Importance	First Stage * Zero vs Partial	First Stage * Full vs Partial	Second Stage * Partial loss only
Junior Ratio(Second Lien Bal/Total Bal)			**Positive**
Migration HELOC's		**Positive**	
Principal Reduction		**Positive**	**Positive**
Lenders Private Insurance(LPI)		**Negative**	
Channel(Broker)			**Positive**
Deficiency			**Positive**
US Retail Bank Loan Indicator			**Negative**
Private Mortgage Insurance(PMI)			**Positive**
First Mortgage Balance			**Positive**
Macroeconomic Attributes			
HPI	**Positive**	**Negative**	**Negative**
Unemployment Rate		**Positive**	**Positive**

***Includes all liens**

For lag and differencing transformations, justify any selections whose lookback periods (lag plus length of differencing window) exceed 12 months.

The Method A RM PD model uses only one lagged attribute specific to the BK01_BK2 transition. This lagged attribute corresponds to - P_M_State_Unemp_Rate_lag12_GOV which is defined as Lagged 12 months state level unemployment rate specifically for government loans. The justification supporting the usage of this attribute is as follows - Higher unemployment rate means lower income of the borrower with similar worsening conditions at the broader macro-economy. Given the pessimistic economic outlook, the borrower has lower ability and willingness to make their mortgage payment. Government loans have a delayed response to unemployment rate because government loans are granted to borrowers with relatively lower debt ratio, which acts as a buffer against immediate delinquency when unemployment rate rises.

If splines/bins or other transformation were used to linearize or discretize variables, provide evidence that model performance improved because of these variables compared to their simpler linear forms, and that the performance improvement was large enough to justify the added complexity.

The Method A RM PD model uses several splines and other transformations as deemed necessary in the scope of model development. For details and justification supporting the spline transformations, please refer to Attachment – '5.1.4 Bubble Plot Justification.xlsx'. For justification around other variable transformations, please refer to attachment 5.1.4 Residential Mortgage PD Specifications file.

If applicable, describe steps taken to prevent inclusion of non-intuitive variables generated by automated statistical/machine learning variable selection algorithms (e.g. Random Forest).

Identify any variables that remain in the final model specification that were generated by automated statistical/machine learning algorithms and demonstrate the mechanism by which their use can be justified on business fundamentals.

As iterated in Section 5.1.4 of the MDD, all candidate pool of attributes were tested on the following two aspects as part of initial variable selection-

1. Business Intuition / Economic Theory: Variables should make sense and there should be an economic rationale for their use in residential mortgage loan modeling.
2. Practical Aspects: Variables should (a) have sufficient historical data availability and (b) be readily available. (c) be well known. Forecasts may not be available for dynamic characteristic variables such as bankruptcy or automatic payment indicators.

After the initial compilation of the raw variables, a stepwise methodology was followed that was based on segmentation, univariate, bivariate, correlation and iterative regression analyses to arrive at the finalized model specification. All rules or logics applied during this variable selection process complied with the Model Testing Guidance. For additional details around the process and executed steps, please refer to Section 5.1.4 of the MDD.

1. Some of the final variable show counter-intuitive sign. For example, in BUK01 - BUK2 transition the sign of FiCO_SP720 is positive, indicating that the higher FICO the higher chance to roll from current to 30-50 DPD. Model Sponsor needs to check business intuition between dependent and independent variables for all final model specification and for all the transition equations.

   Answer – CAMU would like to clarify that there is NO inconsistent or non-monotonic sign issue in the parameter specifications listing if noting the variable explanations/definitions column. There is a main FICO effect with parameter= 'S_M_FicoRefresh' is -0.0104; for FICO between 640 and 720, the parameter effect is -0.0205 = -0.0104 +-0.0101; for FICO > 720, the parameter effect is -0.0113 = -0.0104 +-0.0101 + 0.0092, still being negative. This is essentially saying that above 720, the incremental risk reduction with a FICO increase is decreasing although the relationship that higher the FICO, the lower the propensity to roll to worse, still remains true.

Model	Variable Definition	BK01_BK2
S_M_FicoRefresh	Refreshed FICO at observation month	-0.0104
S_M_FicoRefresh_SP640	Refreshed FICO at observation month, spline flooring at 640	-0.0101
S_M_FicoRefresh_SP720	Refreshed FICO at observation month, spline flooring at 720	0.0092

LGD Model Clarifications

1. LGD Model – Sponsor to provide clear documentation of full losses, partial losses and zero losses.

Answer – The severity model defines the full, partial and zero losses in the following way -

a. Full loss=(Loss_rate>0.97)
b. Zero loss =(Loss_rate<=0)
c. Partial loss=(0<loss_rate<=0.97)

2. LGD Model – Sponsor to provide rationale on how negative losses and loss severity greater than one is treated in the LGD model.

Answer – The following itemized treatments are meted to negative losses and loss severities greater than one-

Negative loss: This is floored at zero, which is a conservative assumption

Zero loss ind=1	N	PRIN_BAL	cum6m_loss
1st	**2,560**	**242,618,390**	**(25,429,134)**
Zero	1,045	92,358,351	-
Negative	1,515	150,260,039	(25,429,134)
2nd	**3,563**	**241,670,193**	**(87,296,065)**
Zero	1,585	113,557,164	-
Negative	1,978	128,113,029	(87,296,065)
Grand Total	**6,123**	**484,288,583**	**(112,725,198)**

Loss severity greater than one: loans with loss rate>0.97 are identified as full loss loan, and will use loss rate=1 (medium value of full loss loans) as forecast loss rate

Analysis Variable : loss_rate																
Portfolio	Full_loss_Ind	N Obs	N	N Miss	Minimum	1st Pctl	5th Pctl	10th Pctl	25th Pctl	50th Pctl	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Maximum	Std Dev
1st	0	57,751	57,751	0.00	0.00	0.00	0.00	0.10	0.27	0.45	0.63	0.79	0.88	0.96	0.97	0.25
	1	15,576	15,576	0.00	0.97	0.97	0.98	0.99	1.00	1.00	1.00	1.12	1.50	2.84	3369.03	41.20
2nd	0	51,199	51,199	0.00	0.00	0.00	0.00	0.20	0.77	0.90	0.94	0.95	0.96	0.97	0.97	0.29
	1	171,514	171,514	0.00	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.02	1.05	1.80	9878.35	23.88

3. LGD Model – The model sponsor states that " Severity is defined as cumulative lifetime loss (up to 6-month after IVP date) divided by Balance at IVP," Model Sponsor needs to provide rationale whether 6-month period is sufficient to capture the loss. Sponsor to provide rationale if 6-month period used in severity calculation is sufficient to capture the loss. How does the loss severity definition consider loss trigger event- first 180 DPD and IVP as loss trigger events? Model Sponsor needs to provide complete documentation of the same

Answer –

A. Cum6m_loss= use 6-month period to capture lifetime loss
B. Total_loss=use loss data up to end of development data (201712) to capture lifetime loss

C. Cum6m_loss is higher than Total_loss, and therefore 6 month is sufficient to capture the lifetime loss, and is conservative.

IVP date	N	cum6m_loss	total_loss	cum6m To total loss pct
2008	46,909	3,248,015,517	3,226,061,058	101%
2009	71,760	5,358,045,384	5,213,073,857	103%
2010	59,704	4,201,298,533	4,096,731,845	103%
2011	46,174	3,241,040,990	3,159,253,297	103%
2012	30,692	2,023,335,341	1,972,762,356	103%
2013	17,068	994,901,208	957,153,593	104%
2014	10,395	544,292,468	523,940,974	104%
2015	7,095	363,553,833	349,746,788	104%
2016	4,601	236,894,632	229,327,704	103%
2017	1,957	95,786,289	94,310,099	102%
Grand Total	296,355	20,307,164,195	19,822,361,571	102%

4. LGD Model – Sponsor to provide explanation of exclusion of pre-crisis and crisis data for CPB loans in the base loss calculation
Answer -

A. Pre-crisis period is not used for 1st/2nd lien model, also there is only one CPB loan in pre-crisis period
B. Loss rate of CPB during stress period is much lower than post-crisis and recent period which is counter-intuitive, therefore stress period data is not used for calculation.

CPB	Count	PRIN_BAL	cum6m_loss	loss_rate
1.Pre_crisis	1	1,055,290	176,445	0.17
2.Stress	38	45,382,872	9,187,513	0.20*
3.Post_Stress	223	227,682,864	76,071,349	0.33

Grand Total	**262**	**274,121,026**	**85,435,306**	**0.31**

**Pre_crisis:<200801, Stress: 200801~201012, Post stress: 201101~201706*

5. LGD model - Sponsor to provide documentation of how lifetime loss is distributed over the forecast horizon for each model usage.
   Answer - There is no loss distribution applied. The loss timing was determined by PCO logic. Please refer to the example within Section 5.1.2 of the MDD to that provides a conceptual example to visualize model component results being used in the end-to-end calculation of Net Credit Loss for a single loan.

6. LGD Model – The LGD model documentation does not explain the definition of default used for the model and whether it is consistent with PD model.
   Answer – Default is defined as IVP in the LGD model, and IVP is consistent with PD.

7. LGD Model – It is stated that "the LGD model does not include expenses", MRM requires the Model Sponsor to provide accounting rules used for the loss calculation, rationale for not considering expenses in the calculation along with any business judgement.

   Answer – The rationale is that expenses, amounts paid to third parties for goods and services including taxes and insurance, do not go into the basis of the loan asset. They are booked to operating expense (not revenue) throughout the entire default and REO process. Credit loss should only be used to recognize the loss of the loan asset, UPB plus any unamortized origination cost, premium or discount.

8. LGD Model –Consideration for any foreclosure and liquidation timeline in estimating the loss severity? Would the loss and recovery be applied some discounts?
   Answer - The LGD model does not have input variables called foreclosure timeline or liquidation timeline. However, the model used state or judicial vs. non-judicial information to account for the loss difference due to state legislation difference around mortgage foreclosures.

9. LGD Model – Explanation of the rationale to exclude pre-crisis and crisis period data for CPB loan base loss calculation, loss rate difference by principal balance and LTV, usage of 300K as cutoff for principal balance, usage of 100 and 120 as cutoffs for CLTV and provide recent period used for calculating stress multiplier.
   Answer -
   i. Pre-crisis period is not used for 1st/2nd lien model due to data quality issue, also there is only one CPB loan in pre-crisis period
   ii. Loss rate of CPB during stress period is much lower than post-crisis and recent period which is counter-intuitive, therefore stress period data is not used for calculation.

CPB	Count	PRIN_BAL	cum6m_loss	loss_rate
1.Pre_crisis	1	1,055,290	176,445	0.17
2.Stress	38	45,382,872	9,187,513	0.20
3.Post_Stress	223	227,682,864	76,071,349	0.33
Grand Total	**262**	**274,121,026**	**85,435,306**	**0.31**

**Pre_crisis:<200801, Stress: 200801~201012, Post_stress: 201101~201706*

10. The LGD model documentation does not provide clear definition of full losses, partial losses and zero losses. Model Sponsor needs to provide the same. Why full loss is modelled as one of the stage one loss type for 1st lien loans?

    Answer – Definition of the full, partial and zero losses have been described on Section 5.1.2 of the MDD. Full losses are not modeled in Stage 2 model because it adversely affects the model's monotonicity to borrower's LTV. Please refer to Section 3.1 for additional details.

11. Rationale for modelling these transitions. For a few PD and LGD, i.e. Second Lien-Stage 1 - Full Vs Partial, it is observed that 1). The intercept is the most significant factor to the model, which could mask the effect of true risk drivers, 2). the coefficient of intercept is highest among all the variables, which makes the model not be able to differentiate the risk between individual loans. Model Sponsor is required to justify this.

    Answer - The test of the intercept in the procedure output tests whether this parameter is equal to zero. For a logistic model, zero intercept implies that the event probability is 0.5. This is a very strong assumption and is most likely not the case in PD and LGD stage 1 models. So, a highly significant intercept in some of these models should not be a problem. The characteristic analysis also shows reasonable risk ranking among major risk drivers.

12. The rationale for modelling these transitions. High R-square is observed for LGD Stage 2 model, 94%, which may indicate potential over-fitting.

    Answer - High R-square not necessary indicates over-fitting. The stability test (table 6.2.2.1) on LGD stage 2 model shows that the model is robust and stable in both In-time out of sample (OOS) and out-of-time sample (long). Therefore, the model is not over-fitting.

13. Was any other modelling approach considered for LGD? What is the rationale for continuing with the two stage approach?

Answer – Yes, alternative modeling approaches were considered for the LGD framework. Please refer to Section 3.2 for details. The rationale for continuing with a two-step approach is that the first step helps identify and segregate the zero/full loss types from the partial losses, and the second step then models the severity of partial losses.

14. Model Sponsor has provided definition of the three types of losses. Please provide rationale on using 0.97 as cutoff for partial loss vs full loss.

Answer - When 1.00 is used as the cutoff point, the "full-loss" loans would be biased because there are quite few loans with loss rate > 1 (see chart below). Thus, a cutoff point of less than 1.00 should be chosen. As shown in the distribution chart below, for the loans with loss rate <1.00, the loan counts decreased quickly in the first 3 bins (highlighted in red dots). Therefore, we used a conservative assumption of 0.97 as the cutoff point.

15. Model Sponsor needs to explain the rationale to exclude pre-crisis and crisis period data for CPB loan base loss calculation, loss rate difference by principal balance and LTV, usage of 300K as cutoff for principal balance, usage of 100 and 120 as cutoffs for CLTV and provide recent period used for calculating stress multiplier. Model Sponsor states that the CPB data contains 1 loan that has loss in pre-crisis period and hence the pre-crisis data is excluded but does not provide justification for the rest of the question. Please provide the same.

Answer - LTV band is used because LTV is the most import driver to determine the loss rate of one loan. Principal balance is used because we found that the relationship of LTV with loss rate is impacted by the balance. At low balance, the relationship between LTV and loss rate is reversed.

We studied the distribution of average loss rate by principal balance groups among the CPB 1st loans within the post-crisis and recent period (i.e., inactive_dt>='01Jan2011'd). As shown in Table 2.1.1, the low balance groups (prin_bal <=300k) don't show the monotonic relationship with loss rate with LTV (highlighted in yellow). Also as shown in Table 2.1.2, at low balance (prin_bal <=300k), the relationship between LTV and loss rate is not monotonic neither (highlighted in yellow). Therefore, 300k is used as the cutoff for principal balance, and LTV band is only applied to the CPB loans with prin_bal>300k.

Table 2.1.1 Loss rate of prin_bal group		
Prin_Bal	Sum of N	Avg Loss_Rate
1: <=100k	13	0.93
2: 100k-200k	18	0.45
3: 200k-300k	14	0.49
4: 300k-500k	42	0.43
6: 500k-800k	41	0.37
7: 800k-1500	53	0.37
8: >1500k	42	0.28
Grand Total	**223**	**0.33**

Table 2.1.2 Loss rate of prin_bal group by LTV		
Prin_Bal	Sum of N	Avg Loss_Rate
1: <=300k	**45**	**0.53**
1: <=100	41	0.55
2: 100-120	2	0.24
3: >120	2	0.56
2: >300k	**178**	**0.33**
1: <=100	119	0.32
2: 100-120	33	0.33
3: >120	26	0.37
Grand Total	**223**	**0.33**

Further,the cross tabulation of loss rate is studied by both the principal balance group and LTV group as shown in Table 2.2. As highlighted in yellow, the lowest 3 LTV groups among the CPB 1st loans with prin_bal >300k do not shown the monotonic increase of the loss rate. So LTV<=100 is grouped together to insure correct monotonic relationship between LTV and loss rate. Considering the sample size and monotonicity of loss rate, LTV>120 is grouped as well. Therefore, the CPB 1st loan's loss rate is grouped (highlighted in green) as Table 2.3.

Table 2.2. Loss rate by LTV group		
LTV group	Sample Size (N)	Avg Loss_Rate
1: <=300k	**45**	**0.53**
2: >300k	**178**	**0.33**
1: <=50	21	0.54
2: 50-80	47	0.25
3: 80-100	51	0.31
4: 100-120	33	0.33
5: 120-150	20	0.35
6: >150	6	0.38
Grand Total	**223**	**0.33**

Table 2.3. Loss Rate by prin_bal and LTV groups		
Group	Sample Size (N)	Avg Loss_Rate
1: <=300k	**45**	**0.53**
2: >300k	**178**	**0.33**
1: <=100	119	0.32
2: 100-120	33	0.33
3: >120	26	0.37
Grand Total	**223**	**0.33**

The stress multiplier is calculated using recent period (201507~201706) *Vs* stress period 200901~201112

16. LGD model shows low sensitivity to stress for second lien (8%) and breaches minimum threshold of 20%. Model Sponsor needs to provide rationale to explain this breach. This is not provided. Model sponsor is requested to provide the same.

Answer – For the second lien loans of the 201806+27m snapshot, the average loss rate of base scenario is already as high as 87%. Hence it is not surprising that this extremely high loss rate of the second liens would be insensitive to macro-economic drivers. With the sensitivity of 8.2%, the average loss rate of our model in the stress scenario is 94%, which is close to the actual loss rate (95%) of stress period (200801-27m).

17. LGD model- The LGD model documentation does not explain how the loss severity definition considers event- first 180 DPD as loss trigger events. The initial write-down may occur upon the first 180 DPD event. Model Sponsor needs to provide corresponding rationale

Answer - Please refer to the attachment – '5.1.2 NCL_Walkthrough' for explanation regarding how the 180DPD loss triggers is captured through the PCO logic calculation. Additional details explained in PCO document can be found in Section 5.1.4 under 'Walkthrough of Model Implementation and NCL (end-to-end) Loss Calculation'.

18. LGD model predicts loss, would the loss and recovery be applied some discounts?

Answer – Please note no discounts are applied for the loss/recovery calculations. Please refer to the attachment – '5.1.2 NCL_Walkthrough' for all pertinent details on the end-to-end loss calculations.

19. MRM Question The MDDT provides data exclusion waterfall (P113) for combined 1st and 2nd lien. Sponsor need to provide data exclusion specific to 2nd lien.

Answer - During LGD model development, the exclusions were applied on overall IVP population but not based on lien position. There is no exclusion which is specific on 2nd lien population. Therefore no data exclusion waterfall is provided for 2nd lien only. The LGD development data is divided into 1st Vs 2nd lien segments after all exclusions have been applied.

Exposure at Default (EAD) Clarifications

1. The model sponsor states that "loan curtailment would be calculated using the formula leveraged by the Capital Markets Prepayment Model to be consistent with the business usages. Curtailment Rate = (Actual Payment amount – Schedule payment amount)/Unpaid Principal balance". Model Sponsor needs to provide rationale on the usage of this formula along with any supporting evidence and explain if any updates/changes are made. 2nd MEA: This information is not provided. Model sponsor is requested to confirm that they have removed curtailment rate.

Answer – Please note that snap-date curtailment is included within the scope of model development. However, the model is not designed to predict future curtailments. Please see Section 5.4 for details. Please see attachment – 5.1.4 Clarify WriteDown Process and Loss Recognition Timing' for details on PCO Chargeoff logic.

2. It is not clear how the curtailment factor is being used in EAD calculations. Model Sponsor is required to show how this is captured in the pseudo example. 2nd MEA: This information is not provided. Model sponsor is requested to confirm that they have removed curtailment rate

Answer – Please note that snap-date curtailment is included within the scope of model development. However, the model is not designed to predict future curtailments. Please see Section 5.4 for details.

Please see attachment – 5.1.4 Clarify WriteDown Process and Loss Recognition Timing' for details on PCO Chargeoff logic

PCO Logic Clarification -

1. Model Sponsor needs to clearly explain the PCO logic including how it is used in LGD model/loss forecasting process, source/owner of PCO logic whether it is external (Government/GSE) or internal (Bank's Finance, Accounting etc.)

Answer - CAMU has conducted rigorous review process with Accounting Team and Risk Management partners responsible for monitoring of Charge-Off policy during 2016 model redevelopment. Further, in order to adapt the most updated policy in PCO Logic, CAMU reinvestigates the PCO logic on an annual basis. This allows the PCO logic to be enhanced as needed while reflecting all applicable policy changes timely. Please refer to the attachment 5.1.4 RE Clarify Write Down Process and Loss Recognition Timing for additional details.

2. Please provide the evidence of approval of the PCO logic from appropriate source.

Answer - Subject to BankMortgage's current Loan Charge-Off Policy, write-down begins before final disposition under many circumstances. In 2013, CAMU implemented a basic PCO logic in the end-to-end loss forecasting process for CMI Transition models to align it with the actual charge-off process used in BAU operations. This document mainly discusses the implementation of PCO logic in general and enhancements of the updated logic. Please refer to the attachment below for additional details.

3. The PCO logic should clearly mention how the losses are distributed across the forecast horizon.

Answer - The Partial Charge-Off (PCO) logic is the key component of the end-to-end loss forecasting process. It determines the timing of losses, i.e. how life time losses are distributed over time. The chart below shows the existing end-to-end loss forecasting process. Please refer to the attachment '5.1.4 PartialChargeOff Logic Document.doc' within the attachments folder for additional details.

4. PCO logic – Sponsor to provide clear explanation of how PCO logic is used in LGD model along with source, evidence of approval, detailed calculation, distribution over time during stress and non-stress periods.

Answer - Refer to attachment 1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY.docx

DV Logic Clarifications

Model Sponsor needs to provide explanation of the DV approach including:

1. Step by step process of haircut estimation

Answer – Please refer to Section 5.1.2 for details around the step by step process of haircut estimation.

2. Rationale and analysis to justify the segmentation into loans prior distressed valuation (PDV) and loans without prior distressed valuation (NPDV);

Answer – All rationales to justify the segmentation into loans prior distressed valuation (PDV) and loans without prior distressed valuation (NPDV) has been discussed in Sections 3.2 and 5.1.3 of the MDD.

3. How haircut is differentiated between baseline and stress scenarios.

Answer - Please refer to Section 5.1.2 for details around the step by step process of haircut estimation and how it is differentiated across base(recent) and stress scenarios.

4. Rationale for using 0% haircut for the base scenario and 10% for the stress scenario for the PDV segment

Answer - Please refer to Section 5.1.2 for details around the step by step process of haircut estimation and how it is differentiated across base(recent) and stress scenarios.

5. Various approaches considered for distressed valuation along with evidence on preferring this approach over others.

Answer – The main alternative to the 2019 DV Logic is the prior year's DV model. Please see discussion in Section 3.2 for additional details.

6. How other dimensions are discarded and combination of occupancy and lien type is considered "significant" for NPDV segment? Sponsor is requested to explain the process how other dimensions were found not "significant" in comparison to "occupancy type and lien position".

Answer – The decision of dimensions/segments included in haircut rate construction is based on both empirical data analysis and economic intuition. Model development conducted haircut rate analysis on various dimensions including occupancy type, lien position, Business, Channel, Property Type, Age of loan, LTV, Original balance, modification, judicial, loan type(FHA vs VA vs Conventional), PMI(Private Mortgage Insurance), Product Type(ARM vs Fixed vs Home Equity. Their differences are shown below It is observed that the haircut rate is mostly different across lien position and occupancy status, which are consistent with the economic intuition. First lien impairment is smaller when the property is under stress. Owner occupied property impairment is also smaller as the property is occupied by owner and used as primary residence, whereas the impairment is bigger if it is used as non-primary residence. We observed little to no difference in other dimensions commonly used for this type of analysis. Some observed difference can still be attributed to lien position and occupancy status, for example, HELOC haircut ratio is significantly lower than ARM/FXD because HELOC are dominated by second lien.

Occupancy	% of total	1-Haircut
1.OwnerOcc	94.46%	0.816
2.NotOwnerOcc	5.54%	0.732

LienPosition	% of total	1-Haircut
1.1stLien	91.68%	0.791
2.2ndLien	8.32%	0.869

Business	% of total	1-Haircut
1.CPB	0.15%	0.800
2.Non-CPB	99.85%	0.805

Channel	% of total	1-Haircut
1.BROK	25.09%	0.796
2.CORR	25.09%	0.785
3.RETL	17.39%	0.836
4.WLST	14.05%	0.815

Property Type	% of total	1-Haircut
1.SFR	88%	0.821
2.CONDOCP	7%	0.797
3.MULTI	2%	0.712
4.OTHER	2%	0.759

Age	% of total	1-Haircut
1.<=2 Yr	2%	0.772
2.(2,3]Yr	6%	0.783
3.(3,4]Yr	11%	0.798
4.(4,5]Yr	12%	0.797

5.(5,6]Yr	12%	0.810
6.(6,10]Yr	38%	0.816
7.>10Yr	20%	0.804

LTV	% of total	1-Haircut
1.<=80	29%	0.796
2.(80,100]	30%	0.803
3.(100,120]	23%	0.817
4.>120	18%	0.819

OrigBal	% of total	1-Haircut
1.OrigBal<=60K	12%	0.816
2.OrigBal(60K,120K]	28%	0.797
3.OrigBal(120K,240K]	37%	0.813
4.OrigBal(240K,800K]	23%	0.816
5.OrigBal800K+	1%	0.757

Modification	% of total	1-Haircut
1.NeverModified	75%	0.797
2.Modified	25%	0.828

Judicial	% of total	1-Haircut
1.Judicial	51%	0.795
2.NonJudicial	49%	0.816

LoanType	% of total	1-Haircut

1.FHA	15%	0.792
2.VA	4%	0.850
3.Conventional	82%	0.805

PMI	% of total	1-Haircut
1.w/ PMI	13.59%	0.797
2.w/o PMI	86.41%	0.806

ProdType	% of total	1-Haircut
1.ARM	14.03%	0.792
2.FXD	83.70%	0.801
3.HLC	2.26%	0.873

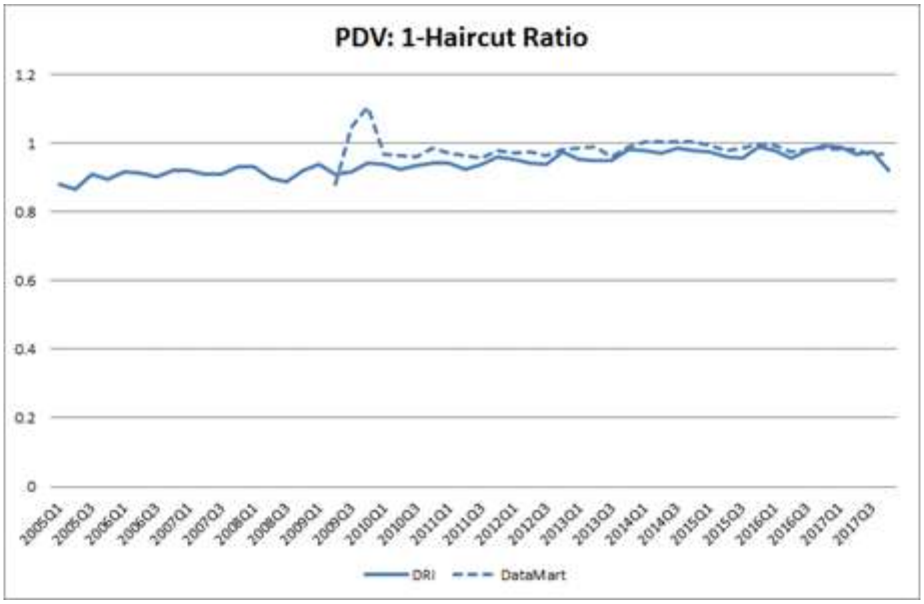
7. The Method A 1st lien model is used for estimating losses from various portfolio segments (e.g. CMI, CPB etc.). Model Sponsor needs to provide rationale on whether the data used as input to DV logic covers these various portfolio segments.2nd MEA: This is not provided. Model sponsor is requested to provide the same.

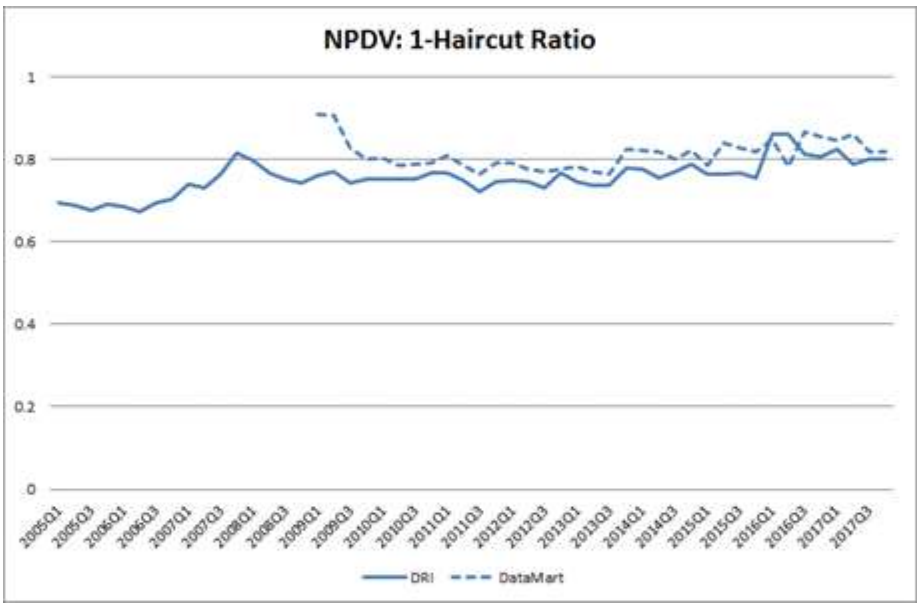
Answer- The entire portfolio – CMI + CPB along with all liens were used in the scope of DV Logic calculation.

8. How haircut is differentiated between baseline and stress scenarios.2nd MEA: This is not provided. Model sponsor has not provided the same. rationale for using 0% haircut for the base scenario and 10% for the stress scenario for the PDV segment2nd MEA: This is not provided. Model sponsor has not provided the same.

Answer - First of all, we are using Risk Datamart data to determine the haircut ratio to be used in the 2019 CCAR model. But Risk datamart data is only available since 2010. Haircut ratio associated with stress period/scenarios were determined with reference to a legacy data source (DRI).

(1-Haircut ratio) based on the DRI data starting from Jan 2005 for the NPDV and PDV segments respectively are given below, overlaying with the ratio derived based on Risk Datamart data.





Following conclusions were made based on the above charts:

1). For the NPDV segment, haircut ratio is relatively stable cross time. Hence the model assigns the same haircut ratio for stress period/scenario as for base period/scenario.

2). For the PDV segment, according to the chart above, haircut ratio is about 10% during the stress period, and 0% in recent period.

PD Model Clarifications –

1. Please provide rational and/or supporting analysis to show that key risk drivers in the PD model will remain significant throughout the forecast horizon.

Answer – For each individual model equation, the parameters are estimated on model development data and then compared across in-time(INT) and full(includes both model development and OOT) samples. Variables that survived in the final model equation had to meet the following three criteria.

i. Parameter sign for all attributes, inclusive of all key risk drivers remains the same across development data, INT, and full sample
ii. Parameter estimates remain statistically significant (P-value <= 0.05) across development data and full sample (significance is not required in INT due to very low volume of some INT samples)
iii. Magnitude of parameter estimate do not have larges shift across development data, INT, and full sample

Please review the Model Stability Section 6.2 for additional details.

2. The final model specification exhibit potential overfitting of the model. Model sponsor needs to provide rationale/supporting evidence to explain: Why not leaving enough most recent period data for OOT sample.

Answer – Please refer to Section 4.1.3 for all pertinent details around the selection of the OOT period.

3. The final model specification exhibit potential overfitting of the model. Model sponsor needs to provide rationale/supporting evidence to explain: a) Why not leaving enough most recent period data for OOT sample. b) The justification of using many boolean/indicator variable such as Katrina/Sandy. c) Increasing number of modeled transition equations.

Answer – Please refer to Section 4.1.3 for all pertinent details around the selection of the OOT period. Based on model sponsor leadership requests, effects from natural disasters such as hurricanes were included in the list of risk drivers considered in the 2019 model redevelopment project in the form of indicator variables. Section 5.1.4 clearly specifies the strategy deployed this year to segment between non-modeled vs modeled equations. This strategy/structure had two-fold advantages –

a) First and foremost, it created a logical approach to logically and statistically separate the modeled vs non-modeled transitions
b) Second, it helped to simplify the model's complexity and assisted in justifying the logic behind each and every modeled equation

Please refer to Section 5.1.4 for additional details.

4. Sponsor to provide rationale for including variables with multicollinearity as evidenced by high VIF (>5).

Answer – The 2019 RM model does not include variables with high degree of multicollinearity. The finalized attributes which have a VIF> 5 are usually capturing spline/interaction effects. Using the BUK01->BUK2 as an example, within this equation for more than 30 variables there is no potential over-fitting as many variables are trying to capture the spline effects, specific product related effects such as I/O and the difference between conventional loans and government loans. Considering that the portfolio coverage includes various products (ARM/FRM, IO/non-IO), portfolios (CMI/CPB), underwriting policy changes over time, Government vs Conventional, etc., this specification is actually highly efficient compared to splitting and developing additional separate equations. Second, the high VIFs of some variables are NOT a concern as all VIF>5 only happens to spline variables or dummy variables to identify a product and its specific effects such as I/O dummy and I/O payment shock. They are expected to have high VIF, which is not a concern from the modeling or statistical perspective. For all other variables, the VIF is much lower than MRM specified cut-off value of 5. The following bullets summarizes the reasons/explanations of VIF>5 for all PD equations.

- Various spline variables, such as CLTV<60, CLTV_60_80, CLTV>80
- A dummy variable + its specific sensitivity effects, such as I/O dummy and I/O payment shock

5. Sponsor to provide rationale and supporting analysis of not including key segmentation variables such as CLTV, FICO etc. as was identified by MRM in the past analysis.

Answer - It is not a good approach to use continuous variable such as FICO, CLTV as segmentation variables as it will create non-continuous sensitivity across these variables. For example, if FICO<=700 and FICO>700 is used as segmentation, the model will consider fico =700 to be very different from Fico=701. These type of continuous variables are used as explanatory variables on the right hand side of the equations.

6. The step-by-step variable eliminating/adding process (for example, due to statistical insignificance or business override) from candidate variable pool to final model variables is not provided. MRM requests model sponsor to provide the same. Please provide step-by-step variable eliminating/adding process

Answer – As iterated in Section 5.1.4, the variable selection process was conducted in a iterative fashion, with adhering to the following rules, that justified the retention or exclusion of any given attribute-

- Variable significance and parameter sign
- Multi-collinearity (VIF) < 5
- Consistent with business interpretation/economic intuition
- Robust in terms of model's goodness of fit across in-sample and out-of-sample based on model's In-sample goodness-of-fit
- Parsimonious with reasonable AIC statistics and one-month ahead fit

Presented below is an example detailing the variable selection process for the RM BUK01 ->BUK 2 transition. Each step within this underlying table illustrates the number of variables dropped from this equation based on each specific criterion. The Final Model only retains variables which is 1) Significant; 2) Conforms with business intuition and is stable across different samples.

Variable Selection Steps		RM_buk 01_buk2
ste	Initial varlist	90

p1		
step2	remove macro-economic change variables defined at different time scale (for example, keep income_12m_ratio, remove income_6m_ratio and income_3m_ratio)	83
step4	Adding variables based on bubble plot, and interactive variables	136
step5	Remove variables with wrong sign	92
step6	Remove variables with p-value >0.05	85
step7	Remove variables with high VIF or having stability issue (changing sign or insignificant on validation sample)	74
step9	Remove variables with low chi-sq	59
step10	Remove variables at final model tuning stage	45

To summarize all PD model equations have been validated across time, based on all the criteria listed above for both model DEV and OOT validation time frames and the justification for every variable in every model has been provided in the attachment' 5.1.4 Residential Mortgage_PD_Specifications.xlsx'.

7. Model sponsor needs to provide explanation of how refreshed FICO and past delinquency variables will be forecasted for use in Model implementation.

Answer - The model uses Refresh FICO only in the BUK01->BUK2, BUK01 -> VP model, i.e. within the current loan models. In the multi-month forecast, Refresh FICO as of snapshot are used. CAMU intentionally did not use Refresh FICO in delinquent equations because we cannot forecast the post snapshot date FICO once a loan becomes delinquent. All delinquency history variables at month t were calculated based on the forecasted probability of delinquency status at month t, and then were used to forecast the delinquency status at month t+1. Currently the model uses life time Ever_60P (B_M_Ever_60P_B1M) dummy indicator in the BUK2->BUK01 and BUK2->BUK3 transition. In the implementation, if a loan was never 60DPD+ at snapshot date, and at month 1, the model forecasted 10% of probability of rolling in to BUK3 and above(60 DPD+), then the Ever_60P at month 1 became 0.1. And it is never 60Plus portion at month 1 becomes 0.9. At month 2, if among the 0.9 never 60DPD portion in the previous month, there is another 20% probability of going to 60DPD+, then new EVER_60P at month 2 equals to 0.9*0.2=0.18. In this case, at the end of month 2, life time Ever_60P becomes 0.1+0.18=0.28. This process is repeated in the multi-month forecast.

8. MRM observed that beginning states were clubbed together (e.g. 0 DPD and 1-29 DPD), Model Sponsor should provide detailed explanation and analysis for grouping the states. How does clubbing/non clubbing affect the model performance?

Answer – As part of the MIS reporting, the business reports most of the RM loans using MBA delinquency status. Within the RM portfolio, most of accounts has a due day at the first day of the month. These loans, if delinquent, will skip the 1~29DPD bucket but directly go to 30~59 DPD (BUK2) in as per the MBA Delinquency Status definition. As a result, it is not uncommon to witness many RM delinquent loans skip the 1~29 DPD bucket. Hence, CAMU considered it prudent to combine 1~29 DPD status with 0 DPD as 1~29DPD is not an accurate reflection of the RM delinquency status.

9. PD Model - Sponsor to provide rationale and evidence for choosing 1% cutoff modeled transition selection and also provide details on the trend of transitioning from deep delinquency to VP by snapshot month. Sponsor to provide details of the methodology used for calculation of roll rates for non-modeled transitions for stress and non-stress period with sufficient rationale.

Answer - Transitions that contribute less than 1% to the starting (or source) bucket as well as less than 1% to its ending (or destination) status are determined as non-modeled. Take BUK4->VP as an example. Based on entire RM data, 0.6% of the loans in bucket 4 at this month will transit to VP at next month. Meanwhile, among all loans ending in VP, only 0.5% comes from bucket 4. Since its contribution to both starting and ending status is less than 1%, BUK4->VP will be determined as a non-modeled.

Take BUK7->VP as another example. 0.6% of current bucket 7 loans will roll to VP next month, however, 3% out of loans ending in VP stems from bucket 7. Therefore, BUK7->VP will be modeled. Detailed evidence for each non-modeled transition can be found under Logic 4 in sheet 'Determine Model vs Non Model' in attachment '5.1.4 Residential Mortgage Modeled vs Non-Modeled Transitions.xlsx'.

CAMU chose 1% as the cutoff in order to ensure there are sufficient number of modeled transitions while at the same time being cognizant of incorporating and subsequently modeling too many volatile transitions. To add some context, if 0.5% was chosen as the cutoff, a few more rare and volatile transitions would have been modeled (for example - BUK6->BUK3). In contrast, choosing a higher cutoff rate such as 2% would have led to the loss of several intuitive transitions which have sufficient number of observations to build a statistical model(for example - BUK3->VP). Hence the 1% cutoff rate was considered prudent in terms of balancing the sufficiency of observations versus the volatility of rare transitions, both of which impact the build of a robust model.

Figure A illustrated below shows the monthly transition rate from deep delinquency to VP while Figure B shows the monthly contribution rate of Deep DLQ to VP.

Details of the methodology used for calculation of non-model roll rates are provided in sheet NonModelRollRate(CONV), NonModelRollRate(GOV) and NonModelRollRate(Trial) in attachment 5.1.4 Residential Mortgage Modeled vs Non-Modeled Transitions.xlsx.

Figure A – Transition Rate – Deep DLQ→ VP

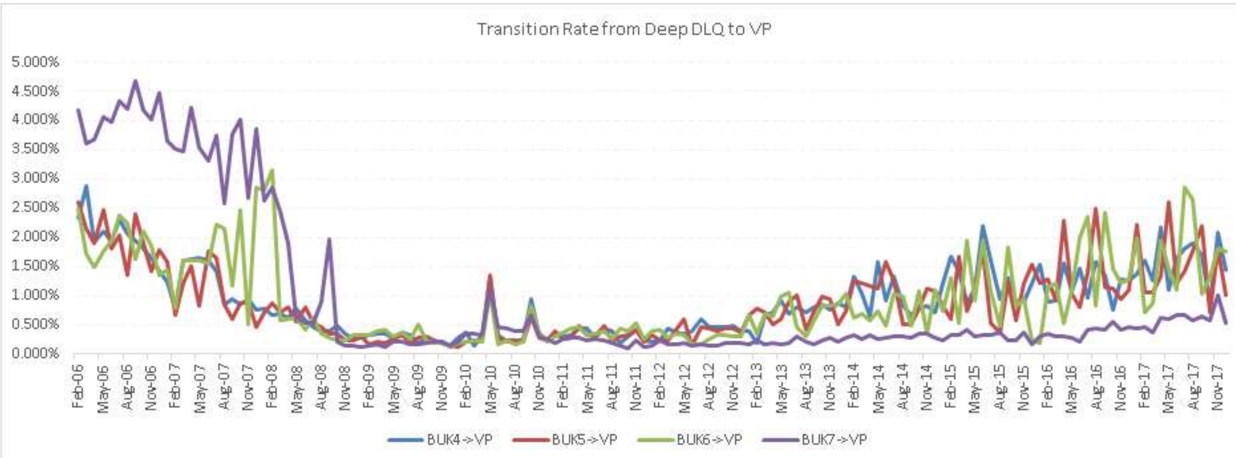
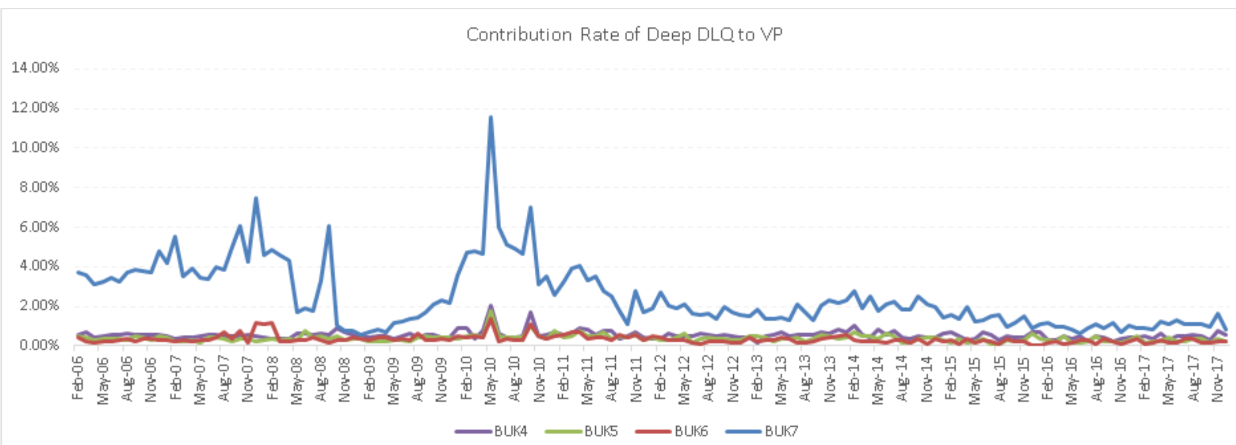


Figure B– Contribution Rate – Deep DLQ→ VP



10. PD Model - The One month ahead analysis shows that the actual transition rates of BUK4->VP are below 1% from 2006 and 2017, and the actual transition rates for BUK7->VP are below 1% from 2008 to 2017. Please explain how the "1%" has been used to determine the modelled transitions.

Answer – Please see response to RM MEA Question # 9 above.

11. PD Model - Model Sponsor needs to provide rationale on how eight transitions that were non-modeled equation in the previous were determined to be modeled equations including business reasoning and also how it was decided to model rare events like BUK7->BUK01 and BUK7->VP. Government and in-trial loans are segmented into three categories. Model Sponsor needs to provide rationale on process of roll rate calculation for these three categories along with any differences between stress and non-stress periods

Answer – Please refer to Section 6.1.9 of the MDD which illustrates the one-month ahead analysis for the non-modeled transitions (provided at segment level).

12. PD Model - Sponsor to explain why additional segmentation was not considered for current (BUK01) loans as they constitute 87% of RM development data. Wrapper: Table 4.1.3.6.1 shows that current loans (BUK01) accounts for 87% of the RM development data (Page 97), have the model sponsor considered any segmentation within BUK01? Also, please note that Method B models has segmented the current loans.

Answer – The 2018 Method B model suite used simulation approach which given its approach had more flexibility in using historical DLQ variables which is why Method B had segmented BUK01 into clean and dirty population. Although Method A did not introduce additional equations, interaction variables have been used to account for segment level differences such as CMI/CPB, GOV/CONV, sensitivity and performance differences across difference origination channels.

13. PD Model - Please provide rationale on PD non-modeled transitions w.r.t.: 1) Using CMI parameters on CPB loans. 2) Government and in-trial loans are segmented into three categories. Model Sponsor needs to provide rationale on process of roll rate calculation for these three categories along with any differences between stress and non-stress periods, sponsor to provide rationale on increase in number of modeled transitions especially for higher delinquency buckets as it may lead to overfitting .

Answer – Please see responses above

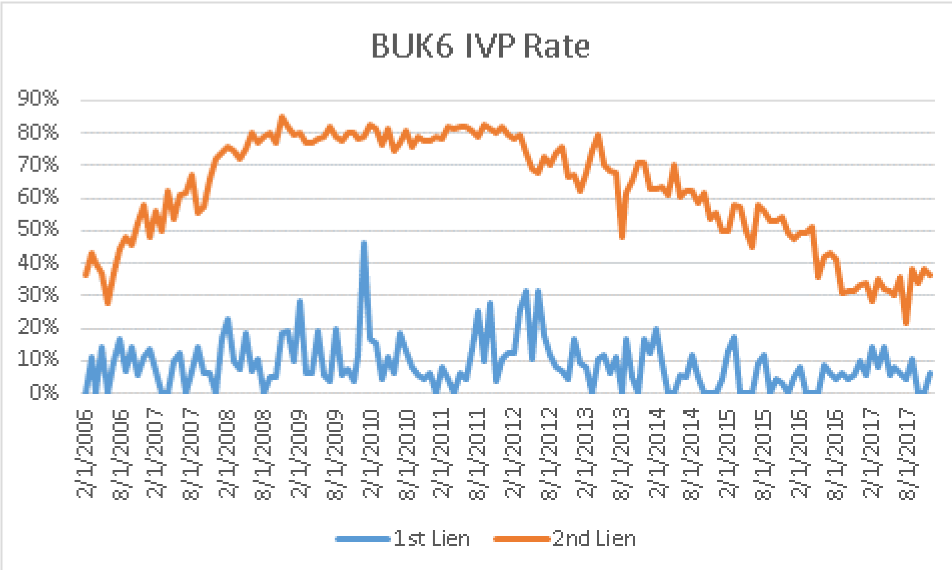
14. PD Model - The VIF of several final variables is greater than 5 (for e.g. transition BUK6 to IVP , variables - B_S_lien2nd_IND and HPI_tm12_ratio_SP1_2nd have VIP of 2295.1 and 2202.9 respectively). Sponsor needs to provide the rationale for considering these variables in the model.

Answer- B_S_lien2nd_IND is used in the BUK6->IVP model because most of the RM 2nd liens are charged off at BUK6. The charge-off rate for second liens is much higher relative to first liens at BUK6, as shown in Figure C below. Second, as illustrated below, the BUK6 to IVP rate has a correlation with the housing price delta such that lower the HPI, the higher the IVP rate. Hence, CAMU considers it prudent to retain both afore-mentioned terms in the model to

- Distinguish first lien and second lien IVP rate from BUK6
- Adequately capture second lien IVP sensitivity to HPI

As HPI_tm12_ratio_SP1_2nd has a negative sign in the BUK6->IVP equation, if we do not include B_S_lien2nd_IND, the model will forecast higher than actual BUK6->IVP rate for first lien but lower than actual rate for second lien. On the other hand, if the model had only included B_S_lien2nd_IND, then the model will lose its desired sensitivity to HPI.

Figure C – BUK6 → IVP



15l Prepayment burnout is an input variable for model BK1 -> VP, but not in the summarized specification provided. Please correct the specification.

Answer – Please refer to Section 5.1.4 attachment – '5.1.4 Residential Mortgage PD Model Specification with explanations.xlsx' for details on this attribute.

16. The One-month ahead analysis shows that the actual transition rates of BUK4->VP are below 1% from 2006 and 2017, and the actual transition rates for BUK7->VP are below 1% from 2008 to 2017. Please explain how the "1%" has been used to determine the modelled transitions

Answer - Please refer to responses to Question # 9 above.

17. The one month ahead analysis shows that the prediction error tends to be large in stress periods for a few transitions, such as BUK2->BUK4, BUK2->BUK5, BUK2->BUK6, BUK2->BUK7, BUK2->IVP, BUK3->IVP, BUK4->IVP, BUK5->IVP. Please explain the rationale of large prediction errors during stress periods.

Answer – The methodology for calculating the non-modeled transitions rates have been overhauled from its prior version (submitted 10/01/2018). Please refer to the Section 6.1.9 which has the updated one-month ahead analyses for the non-modeled transitions.

18. Please provide rationale for large shift of model parameter estimation in a few equations of the PD and LGD model. Please provide the rationale for the same. E.g., B_M_IN_TRIAL_IND in BUK01 to VP transition, and R_M_State_UnempB12M_SP1_MOBGT120 in BUK3 to BUK01 transition

Answer - The B_M_In_TRIAL_IND is used in the BUK01->VP transitions to capture the effect that loans in trial have had experienced in the past some form of financial hardship as such these loans are less likely to VP. The effect was kept in the model to separate the significant difference between trial versus non-trial loans. Within this resubmitted version of the RM MDD (resubmission – 11/1/2018), R_M_State_UnempB12M_SP1_MOBGT120 in BUK3 to BUK01 has been removed. Please refer to the updated – '5.1.4 Residential Mortgage PD Model Specification .xlsx' file for details.

19. Some of the transitions are observed to have too many variables (for e.g. BUK 1 to BUK2 transition 1 has 46 variables) which may lead to overfitting of the model. Please justify the usage of these variables in terms of marginal contribution, such as incremental improvement in C-statistics, of each final variable and explain the impact of each variable. Further MRM question in 2nd MEA: This information is not provided. The Model Sponsor is requested to provide justification on incremental benefit for each variable.

Answer – As reiterated in Section 5.1.4 and Response to MRM MEA Question # 6 above, CAMU utilizes the following set of rules to justify the retention/exclusion of any given attribute from any modeled equation. As such, C-statistics was not the only criterion that was used. Further, please note many attributes within the final model specifications capture the segment level performance that actually greatly aids in keeping the model structure simplified, while still being able to track the differentiated sub-segment level performance.

- Variable significance and parameter sign
- Multi-collinearity (VIF) < 5
- Consistent with business interpretation/economic intuition
- Robust in terms of model's goodness of fit across in-sample and out-of-sample based on model's In-sample goodness-of-fit
- Parsimonious with reasonable AIC statistics and one-month ahead fit

20. MRM observed that the model contains many Boolean/indicator variables such as Katrina, Sandy etc. Has the developer checked the performance of the model without these variables as well as justified the usage of these variables in the model? 2nd MEA: This question is partially addressed.

o The Model Sponsor responded that the natural disaster variables are included as risk drivers based on Model Sponsor leadership requests. Please provide rationale and supporting evidence. Also, share details on how these variables are created.

o Model Sponsor has not provided rationale for usage of all other Boolean/Indicator variables. Please provide the same.

Answer - Please see attachment "5.1.4 Natural Disaster Information" which has supporting evidence regarding the addition of these attributes in the 2019 model development process, based on senior leadership request. The zipped file also contains information around the creation of these attributes.

21. Model Sponsor is requested to provide the following:

Confirm that Curtailment rate is removed from the model. If yes, then how it is captured in the model.

Describe PCO charge-off logic along with source (internal vs external) and evidence. 2. Please provide the evidence of approval of the PCO logic from appropriate source...

Answer – Please note that snap-date curtailment is included within the scope of model development. However, the model is not designed to predict future curtailments. Please see Section 5.4 for details. Please see attachment – 5.1.4 Clarify WriteDown Process and Loss Recognition Timing' for details on PCO Chargeoff logic.

22. Model Sponsor has mentioned that non-modeled transitions in the previous model are classified as modeled equations this time due to the use of the waterfall approach but they have not provided business justification. Please provide business justification for modeling rate events mentioned above.

Answer – Please see section 5.1.4 for details that were added for this specific MEA.

23. Please explain why this approach was selected as compared to other approaches.

Answer – Please see Section 3.2 for a detailed discussion on the three alternative approaches that were considered.

24. Rationale of the selection of preferred segmentation approach – business decision versus statistical reasoning

Answer - The model's finalized segmentation aligned with common business expectations and were further validated by statistical analyses that included decision tree and bivariate bubble plots. The finalized segments of interest were further validated/fine-tuned based on iterative regression runs. Please refer to Section 5.1.3 for additional details.

25. Model sponsor has employed four logics in determining modeled transition (page 196). Has the model sponsor considered business practice in the development of this logic? 2nd MEA: This is not addressed. Model Sponsor is requested to do the same.

Answer – The waterfall logic has been updated with more clear narrative and explanation . Yes, business practices were considered as part of Logic # 1. Please refer to Section 5.1.4 for more details.

26. The calculation of policy in the attachment '1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY' (Page 8) utilizes Broker Price Opinion (BPO) and a scalar multiplier of 0.9. Model Sponsor needs to provide source of BPO and rationale on how the scalar multiplier was determined.
Answer – Please see attachment 5.1.4 Clarify WriteDown Process and Loss Recognition Timing' for details on the BPO multiplier.

27. Please provide a working excel for an account to visualize the calculation of each of the component and the end to end calculations.2nd MEA: This is not provided. Model sponsor has not provided the same.

Answer – This has been provided under 'Walkthrough of Model Implementation and NCL (end-to-end) Loss Calculation' within Section 5.1.2.

28. Please confirm that the estimates from the 80% DEV data is used as the final model specification.

Answer - Specifically for the 2019 Method A Residential Mortgage PD model, data from Feb 2006 to Dec 2011 and from Apr 2014 to Dec 2017 is considered as in-time sample, of which 80 percent is used as model development data and the remaining 20 percent is used as INT validation data. The data from Jan 2012 to Mar 2014 (27 months) is considered as out-of-time data.

For the 2019 Method A LGD model, data from Jan 2008-to Dec 2011 and from Apr 2014-to Jun 2017 is considered as in-time sample, of which 80% percent is used as model development data and the remaining 20% percent is used as INT validation data. The data from Jan 2012- Mar 2014 (27 months) is considered as out-of-time data.

29. Prepayment burnout is an input variable for model BK1 -> VP, but not in the summarized specification provided. Please correct the specification.

Answer – This has been corrected in the summarized table.

30. Model Sponsor is requested to provide rationale of not following MTG recommended threshold of VIF <5. The Model Sponsor has provided justification for only the two examples. Please provide justification for all instances of variables with high VIF.

Answer - One by one justifications are given below:

BUK01->BUK2:

B_S_BROK_IND	16.1
R_M_State_UnempB12M_BROK_CONV	16.0

The above are broker channel main effect and its interaction with unemployment

B_S_INT_ONLY_IND	83.8
P_AND_I_ratio_IO	65.3
R_M_State_UnempB12M_IO_CONV	18.8

IO main effect and its interactions, to capture 1. different risk levels of IO loans and non-IO loans 2. increased delinquency risk when IO reset happens measured by payment shock 3. special sensitivity of IO loans to YOY unemployment rate change

S_M_FicoRefresh	8.1
S_M_FicoRefresh_SP640	17.9
S_M_FicoRefresh_SP720	6.5

FICO main effect and its splines

MTM_CLTV_SP40_1st_CONV	5.3
MTM_CLTV_SP80_1st_CONV	12.5
MTM_CLTV_SP80_2nd_CONV	5.2

MTM_CLTV_SP40_1st_CONV and MTM_CLTV_SP80_1st_CONV are used to capture the non-linearity relationship between MTM_CLTV and delinquency risk. MTM_CLTV_SP80_1st_CONV and MTM_CLTV_SP80_2nd_CONV are used to capture the different sensitivity to MTM_CLTV of 1st lien and 2nd lien.

MTM_CLTV_SP80_1st_CONV is a spline of MTM_CLTV_SP40_1st_CONV, they overlap on the Conventional 1st lien when CLTV >=80. So there is high VIF associated with them.

Besides, RM modeling data is dominated by 1st lien. MTM_CLTV_SP80_2nd_CONV always equals to 0 on the 1st liens so it is negatively associated with MTM_CLTV_SP40_1st_CONV and MTM_CLTV_SP80_1st_CONV which are always non 0 on 1st liens. So there is also high VIF on MTM_CLTV_SP80_2nd_CONV.

P_M_State_Unemp_Rate_lag12_GOV	5.9
MTM_CLTV_SP80_GOVGT3	5.1

P_M_State_Unemp_Rate_lag12_GOV and MTM_CLTV_SP80_GOVGT3 are used to capture the specific sensivity to YOY unemployment and MTM_CLTV effect by GOV loans.

In BUK01->BUK2 data, more than 90% are conventional loans and MTM_CLTV_SP80_GOVGT3 and P_M_State_Unemp_Rate_lag12_GOV always equal to 0 on these loans. So over the BUK01->BUK2 data, these two variables are positively associated with each other. Their correlation is only 0.2873 within the government loans.

BUK3->BUK01

D_M_PRIN_BAL_in10K_CMI_HPIDecr	36.4
D_M_PRIN_BAL_SP100K_CMI_HPIDecr	34.3

UPB Main effect and its splines

BUK6->IVP

D_M_PRIN_BAL_in10K_CMI_1st	60.8
D_M_PRIN_BAL_SP100K_CMI_1st	61.5

UPB Main effect and its splines

B_S_lien2nd_IND	2316.2
P_M_State_Unemp_Rate_2nd	10.2
HPI_tm12_ratio_SP1_2nd	2204.2
MTM_CLTV_SP40LSP150_2nd	8.8

2nd lien main effect and its interaction with unemployment, MTM_CLTV and hpi_tm12_ratio. On all B_S_lien2nd_IND=0, P_M_State_Unemp_Rate_2nd, HPI_tm12_ratio_SP1_2nd and MTM_CLTV_SP40LSP150_2nd always equal to 0. So they are all positively associated with B_S_Lien2nd_IND. Within the 2nd liens, correlation between P_M_State_Unemp_Rate and HPI_tm12_ratio_sp1 is only -0.25803, between P_M_State_Unemp_Rate and MTM_CLTV_SP40LSP150 is 0.19447, between HPI_tm12_ratio_sp1 and MTM_CLTV_SP40LSP150 is -0.21601.

31. The one month ahead analysis shows that the prediction error tends to be large in stress periods for a few transitions, such as BUK2->BUK4, BUK2->BUK5, BUK2->BUK6, BUK2->BUK7, BUK2->IVP, BUK3->IVP, BUK4->IVP, BUK5->IVP. Please explain the rationale of large prediction errors during stress periods.

Answer - BUK2->BUK4, BUK5, BUK6, BUK7, IVP, and BUK3->IVP are non-modeled transitions. The stress parameters were determined by volume weighted mean of transition rates based on data from Feb 2006 to Dec 2011. In addition, in the case when these transition rate was lower than that based on the recent data from Mar 2014 to Dec 2017, the stress parameter would be set to the recent parameters. BUK4->IVP and BUK5->IVP rate were relatively low in late 2008 to 2009, when the HPI drops to its bottom and unemployment reaches its high, if compared to 2011 and 2012. The middle delinquent IVP is more of a business behavior than a borrower's choice hence the it's difficult for the model to fully capture the historical trend using variables with intuitive signs. On the other hand, the model is able to forecasted the significantly decreased BUK4, BUK5 IVP rate in recent period.

32. Model Robustness and Stability: Please confirm if the stability analysis provided is for DEV vs in time vs Long Sample. The stability excels provided mentions the results for Dev vs In time vs OOT whereas the MDDT shows DEV vs in time vs Long Sample. High model parameter shift were observed for few of the equations of the PD and LGD model. Please provide the rationale for the same. E.g., B_M_IN_TRIAL_IND in BUK01 to VP transition, and R_M_State_UnempB12M_SP1_MOBGT120 in BUK3 to BUK01 transition.

Answer - In the final RM model, four variables have relevant large coefficient changes across DEV sample and INT sample (coef on INT/coef on DEV>2 or <1/2). They are:

BUK3->BUK3	**DEV**	**INT**	**INT/DEV**
Post2010_Orig	0.2952	0.1354	0.459
BUK5->BUK4	**DEV**	**INT**	**INT/DEV**
D_M_PRIN_BAL_LE2K	-1.4189	-0.6967	0.491
BUK5->IVP	**DEV**	**INT**	**INT/DEV**
B_S_pmi_IND	-0.7994	-1.9678	2.462
BUK6->BUK6	**DEV**	**INT**	**INT/DEV**
D_M_PRIN_BAL_in10K_CMI_HPIDecr	-0.00203	-0.00435	2.143

Post2010_Orig was kept in the BUK3->BUK3 model, to capture the better performance of loans originated after the crisis via stringent underwriting standards.

D_M_PRIN_BAL_LE2K was kept in BUK5->BUK4 model, to capture the effect that borrowers with a very small amount of UPB left in their account might stop making payment because they believe that they have already paid down their mortgage completely. They are less likely to cure and more likely to straight roll into deeper delinquent bucket. Such a behavior is seen on all delinquent buckets.

B_S_pmi_IND was kept in BUK5->IVP model. The liquidation process of a loan with PMI usually requires Bank to meet higher standards than the non-PMI loans. So it usually less likely for a loan to be charged off at BUK5.

The above effects included in the model are business intuitive. Relevant large coefficient shift on the above three variables are largely attributable to the low number of observations associated with the particular category in the model in-time validation sample.

D_M_PRIN_BAL_in10K_CMI_HPIDecr was kept in BUK6->BUK6 model. The large coefficient change is because of the low coefficient associated with the attributes.

There is no variable have large coefficient change across DEV sample and Full sample.

33. Data Transformation: For the use of spline transformations, the non-linear trend should be explainable and spline regression should be used only if it is fully justified and intuitive. The non-linearity hampers the risk ranking of log odds in the model. Sponsors/Developers should justify the choice of spline in place of variable banding that would have ensured risk ranking.

Answer - Bubble plots have been added in U:\GROUPS\CAMU\CAMU 2018 Projects\CCAR 2019\Method A\MDDT_working_updates\RM\Attachments\5.1.4 RM_PD_Bubble_Plots.xlsx

34. Model Variable Selection: Government and in-trial loans are segmented into three categories. Model Sponsor needs to provide rationale on process of roll rate calculation for these three categories along with any differences between stress and non-stress period.

the document states "CPB loans would not have their specific parameters" and it uses CMI parameters. Sponsors are required to provide more details on the same

As per the PD model specification in the worksheet 'To Worse (2mdl)' on attachment in section 5.1.4, payment shock variables are not included in two going-worse models (BUK01-BUK2, BUK2-BUK3). Model Sponsor needs to provide rationale on the exclusion of payment shock variables from these two models.

Answer - BUK01->BUK2 has payment shock variables associated with IO. The name of the variable is **P_AND_I_ratio_IO**. Paymentshock variable associated with ARM **P_AND_I_ratio_ARM** didn't have intuitive sign in BUK01->BUK2 model hence was not kept.

P_AND_I_ratio_ARM are not significant in BUK2->BUK3. Chi-sq of **P_AND_I_ratio_IO** in BUK2->BUK3 is less than 10, so it is not kept in the final model for parsimoniousness.

35. As per the PD model specification in the worksheet 'IVP (4mdl)' on attachment in section 5.1.4, CLTV is included in the BUK6->IVP and BUK7-> IVP models, yet excluded in the BUK4->IVP, BUK5->IVP, and BUK7->IVP models. Model Sponsor needs to provide rationale on the exclusion of CLTV from these two models?

Answer - CLTV is insignificant in BUK4->IVP. It has wrong sign in BUK5->IVP. It is in BUK7->IVP model.

36. Transition matrix is a one-month ahead analysis, the model is used to forecast 27 months' loan movement. Model Sponsor is required to justify if the key risk drivers in the PD model will remain significant in the 27 month forecast period?

Answer - Please refer to Section 6.1.9 which supports the discriminatory power and stability of the model. The Method A models constitute a transition framework for both model development and model implementation. The transition framework always leverages month 't' information/ all relevant risk drivers to predict the outcome in month 't +1'. Please note that the transition framework does not use the information in t = 0 to predict 't+1' as commonly done with a hazard modeling framework. As such the approach deployed for Method A is consistent between how the model is developed and how it is implemented for forecasting purposes.

37. For non-modeled transitions: Government and in-trial loans are segmented into three categories. Model Sponsor needs to provide rationale on process of roll rate calculation for these three categories along with any differences between stress and non-stress periods

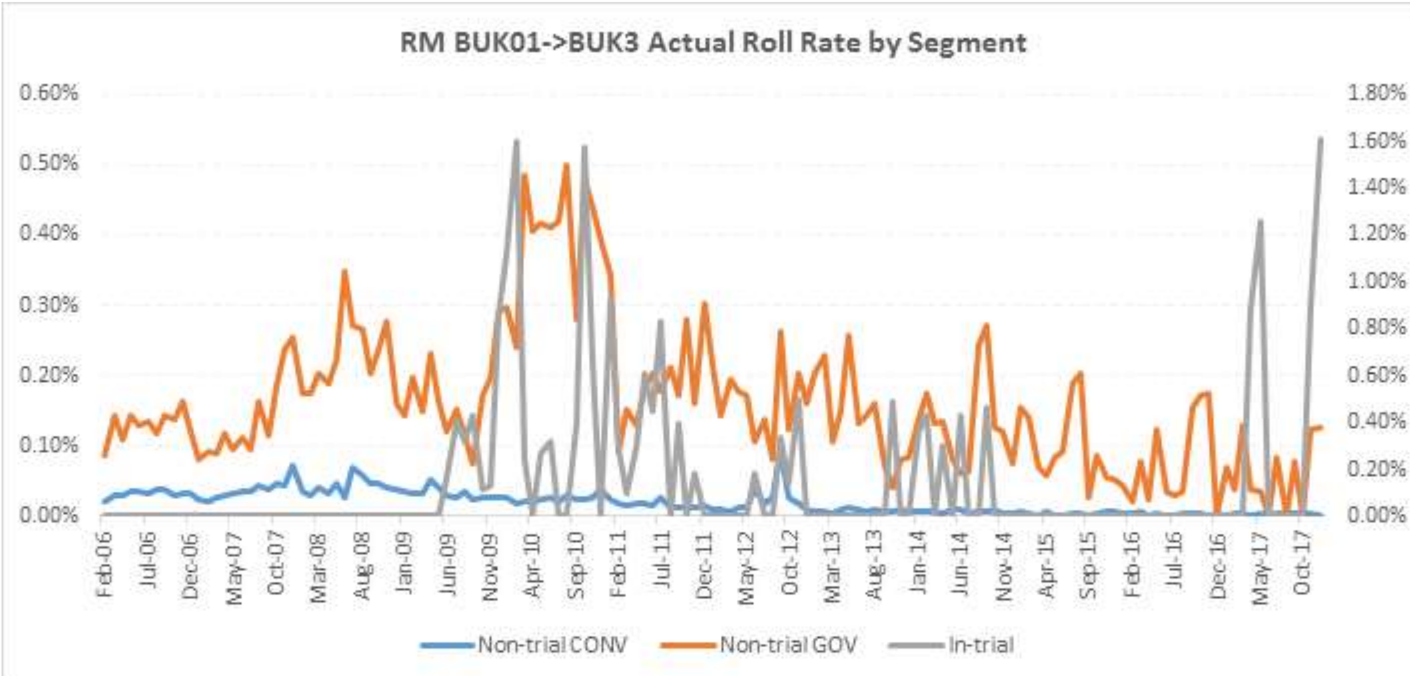
Answer - First, roll rates are observed to be distinguished across the population of government, conventional and in-trial loans.

Second, for each segment, the roll rate is very different for stress and recent time.

With the differences being observed, each non-model is segmented into non-trial conventional, non-trial government and in-trial loans and each segment has a stress roll rate and recent roll rate.

Take BUK01->BUK3 as an example. Figure below shows its actual roll rate by segment. The left axis is for non-trial conventional(COVN) and non-trial government(GOV) loans while the right axis is for in-trial loans. It shows that these three segments have significantly different roll rates across both stress and recent time periods. In 2008, the roll rate for non-trial government loan is about 5 times higher than the conventional loan. In 2017, the roll rate for conventional loan is close to zero while the government loan is higher and more volatile. Trial loans occurred in Apr 2009 and has most volatile roll rate.

Please refer to the attached *5.1.4 NonModel_Actual_RollRate.xlsx* for the actual roll rate by segment for all non-models.



38. For the non-modeled transitions: the document states "CPB loans would not have their specific parameters" and it uses CMI parameters. Sponsors are required to provide more details on the same

Answer - RM has not separated out the empirical lookup non-modeled rate for CMI vs CPB loans. The non-modeled rates were derived based on the population of combined CMI and CPB. Please refer to 5.1.4 for details around RM's non-modeled approach.

39. As per PD model specification in the worksheet 'VP (4mdl)' on attachment in section 5.1.4, interest spread variables are included in the BUK01->VP models, yet excluded in the BUK2->VP, BUK3->VP, and BUK7->VP models. Model Sponsor needs to provide rationale on excluding interest spread variables from the three prepayment models?

Answer - Interest rate spread variable is note rate subtracting swap rate. It measures refinancing incentives. Higher spread rate, more likely to refinance and thus prepay. Prepayment through refinancing is not applicable to a loan in bucket 2 and above since its credit quality has already been impaired. Therefore, interest rate spread variable was only used in the model BUK01->VP.

40. Transition matrix is a one-month ahead analysis, the model is used to forecast 27 months' loan movement. Model Sponsor is required to justify if the key risk drivers in the PD model will remain significant in the 27 month forecast period?

Answer - Please refer to Section 6.1.9 which supports the discriminatory power and stability of the model. The Method A models constitute a transition framework for both model development and model implementation. The transition framework always leverages month 't' information/ all relevant risk drivers to predict the outcome in month 't +1'. Please note that the transition framework does not use the information in t = 0 to predict 't+1' as commonly done with a hazard modeling framework. As such the approach deployed for Method A is consistent between how the model is developed and how it is implemented for forecasting purposes.

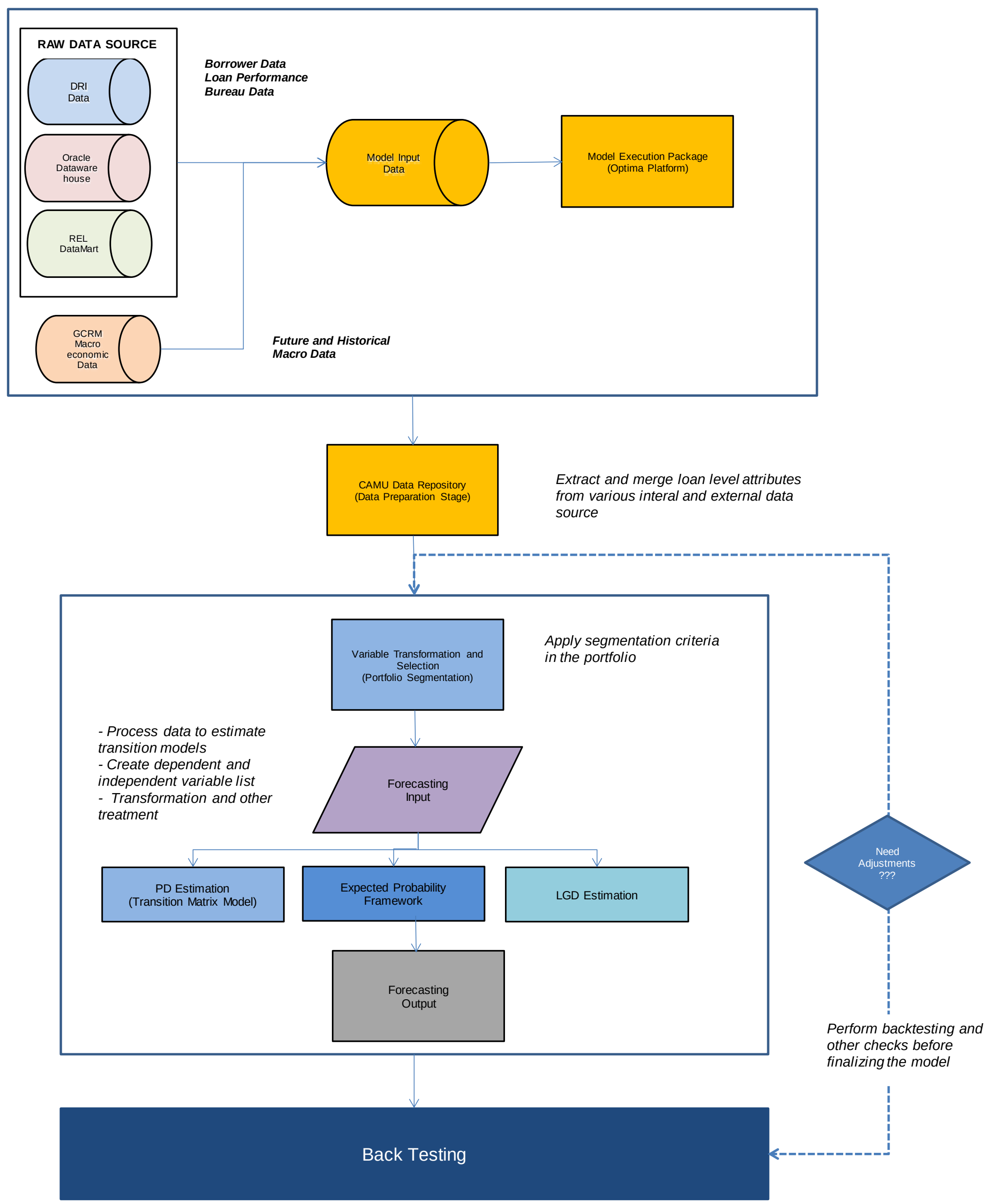
41. Please confirm if the estimates from the 80% dev data is used as the final model?

Answer - As stated at the end of MDDs section 6.2.1, parameter estimated on the 80% model development data is used in the final model.

5.1.5 Model Development Process Flow

[Please provide a flow chart describing the model development process. For example, provide summary process flow chart depicting data sources, model inputs, modeling/estimation components (e.g. PD, LGD, EAD, loss timing curves), model outputs, key business assumptions (e.g. new volume, prepayment).]

Figure 5.1.5 Method A Residential Mortgage Model Process Flow



5.2 Model Specification

This section will provide additional information to the "Automated Document" Section A.3.: "Inputs, Outputs & Assumptions" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

5.2.1 Model Input, Output and Mathematical Equations

[Documentation should include an analysis of the model output, including the congruence of inputs with the assumed economic scenario, the justification of any qualitative adjustment, along with the statistical analysis used to support the model output.

Provide definitions and detailed technical descriptions of model inputs (including dependent variable), model algorithm, and model output (including independent variables). State the final model equation (if applicable). If model nesting is present, feeder models' inputs, algorithms, and outputs must be technically described. A flowchart of the model, showing Input to Algorithm to Output from original model inputs to the final model output(s), must be separately presented.

Example*: For a top-down portfolio segmentation model, individual regression equations are estimated for net charge-off rates. Below is the final form.*

$$Net\ chargeoff\ rate_t = \alpha + \beta_1 * Variable1 + \beta_2 * Variable2 + \cdots + Error$$

The dependent variable must be described in detail. The form of each variable and its time horizon, as it relates to the dependent variable, must be presented as well. The business intuition of the relationship between the dependent variable and the independent variables should also be described in detail.]

For 2019 RM model suite, model inputs included loan-level characteristics as well as regional and national level macro-economic factors including HPI, gross domestic product(GDP), state level income, S&P500 index, unemployment and interest rates. The PD transition model's outputs include unit(s) and balance(s) of delinquencies, voluntary payoff (VP) and involuntary payoff (IVP or default), as well as unit(s) and balance(s) of gross credit losses which are the output of the PD model, incorporating the EAD logic and the newly developed DV Logic. The 2019 LGD model estimates the losses by their outcome type (Full, Partial and Zero losses) respectively. The LGD model inputs include loan-level characteristics as well as regional and national level macro-economic factors as mentioned above. The LGD model outputs include lifetime LGD and lifetime NCL, which is then distributed to monthly NCL based on the forecasting horizon and stated model usages across the applicable scenarios (base, stress, and flat). The detailed model inputs have been described in section 5.1.2 under model design, logic and methodology. For the complete listing of the PD and LGD model input variables, please see the attachments and narrative corresponding to the Variable Selection Process, as illustrated in Section 5.1.4.

5.2.2 Final Model Specification

[Explain the framework that links business drivers to outputs. Consider the business and risk drivers that the approach should incorporate (e.g., macro-variables, currency risk). Provide evidence, economic theory or intuition that explains the relationship.]

- What business/ risk drivers should be captured?

  [Provide a high-level description of the methodology employed and comment on why it is appropriate for this application.]

Managing credit risk for the real estate portfolio requires the ability to forecast aggregate losses on existing loans, predict the length of time that loans will be on the books before prepayment or becoming delinquent or default, analyze the expected performance of particular segments in the existing portfolio, and demonstrate the resiliency of the portfolio to business cycle shocks. Typically a forecasting model with transition structure and nonstationary transition probabilities is used to model the life of a mortgage while logistic and regression models are used to estimate severity of losses. These models are integrated into a system that allows model-end users to depict the expected performance of individual loans and portfolio segments under different economic scenarios. With this information, business and business users can establish appropriate loss reserves, suggest pricing differentials to compensate for risk, and make strategic lending decisions.

The 2019 Method A RM PD Model is a loan-level transition model designed to produce forecasts of delinquency status as well as both voluntary and involuntary termination. With a monthly forecast frequency and a horizon that can be extended throughout the entire life of the loan, this model is designed to work specifically and exclusively for never modified residential mortgage loans in the Bank US consumer mortgage portfolios. It is designed to meet regulatory requirements regarding stress testing as well as serve other purported non-CCAR usages.

The NA Mortgage Net Credit Loss Severity Model estimates the lifetime expected dollar amount of ending net receivables (ENR) that is not collectible due to a mortgage default by the mortgagor. When combined with the PD model and EAD logic of a default incidence model and the DV Logic, the model can produce loan-level loss forecasts.

An enormous literature in credit risk has been fostered by both academics in finance and practitioners in industry. Starting from the publication of Foster and Van Order (1984), Quigley and Van Order (1995), Defranco (2001), and Deng, Quigley, Van Order (2000) and subsequent industry research done by CoreLogic, Loan Performance, Beyond Bond Inc., Moody's Analytics have sufficiently demonstrated that all mortgage credit risk models are influenced by a core set of input variables which eventually affect the payoff, default and delinquency and losses. These variables can be classified as loan level attributes which can be static as of origination or dynamic, i.e. refreshed periodically depending on the evolving portfolio performance to date and the borrowers' financial health; bureau attributes which are collected by the credit reporting bureaus and the broader macroeconomic factors that indicates the economy's state of affairs as well as the market sentiment. Apart from these afore-mentioned risk drivers, natural disasters (Sandy, Katrina, etc.) and periodic seasonality's (increase in home sales before the start of school year, etc.) witnessed in the real estate market also affect the credit worthiness of the homeowner. Based on model sponsor leadership requests, effects from natural disasters such as hurricanes were included in the list of risk drivers considered in the 2019 model redevelopment project in the form of indicator variables. All of these attributes taken together fosters the creation of a credit risk model that can leverage both loan level and macro level data in a single model suite to generate estimates of delinquency default or prepay and the expected losses.

The business / risk drivers in the 2019 RM model suite (both PD and LGD) are classified as follows:

Loan Level Attributes

- Static
  - Property type
  - Loan purpose (Purchase/Refi/Cashout)
  - Co-borrower
  - Occupancy status
  - Loan documentation
  - Property type – Condo vs multiunit vs Single family home
  - Product type – Residential mortgage vs Home equity loan
  - Original interest rate spread
  - Origination vintage
  - Decision FICO
  - Lien
  - Loan type (Conventional vs Government)
  - Business entity (CMI vs CPB)
  - Origination Channel
  - Loan Type – Fixed vs ARM vs IO vs Balloon
  - Prepayment penalty
  - Private mortgage insurance(PMI) indicator
  - Property geography
  - Property state level foreclosure law variables (state specific drivers) such as deficiency, judicial, redemption
- Dynamic
  - Loan age
  - Mark to market LTV / Mark to market CLTV
  - State level foreclosure timeline
  - Refreshed FICO
  - Current note rate
  - Updated unpaid principal balance

- Reset timing – specific to ARM/IO or Home Equity Loans
- Payment shocks post reset
- Modification history
- Delinquency history
- Interest rate spread
- Payment history
- HPI burn-in effect
- Prepayment interest rate burn-out effect
- Foreclosure moratorium

➢ Business Drivers
- In-trial

➢ Bureau Attributes
- Age since oldest trade
- Age of oldest mortgage
- Number of open mortgages

➢ Seasonality Attributes
- Month specific dummy

➢ Natural Disasters
- Sandy
- Katrina

➢ Macro-economic Attributes
- HPI
- Interest rates
- Unemployment rate
- Gross Domestic Product(GDP)
- State level income
- S&P500 index

Since the 2019 Method A RM model suite is a redeveloped model, all of the above variables were used in the candidate pool during the model's initial testing and fit.

The PD model used Jan-2012 to March-2014 period as out-of-time validation sample and retained data between Feb 2006 - Dec 2011 and Apr 2014 -Dec 2017 as in-time sample. The LGD model used Jan-2012 to March-2014 period as out-of-time validation sample and retained data between 01/2008~12/2011 & 04/2014~06/2017 as in-time sample. The reasons for doing so are three-fold –

1. First and foremost, as per limitation # 19546, it was recommended to include the most recent time as part of the model development sample. To comply with MRM's requirements, CAMU chose Jan-2012 to March-2014 period as out-of-time validation sample. The in-time sample was so chosen to retain both the macro-economic stress and the recent timeframes thus incorporating the effect of one full business (and interest rate) cycle.

2. Second, the inclusion of the most recent time helped to factor in the reset risk, primarily for the loans with adjustable rates , especially given the steadily increasing rates since 2016Q1.

3. Third, the inclusion of the recent period data helped in capturing the go-forward state of the business in terms of origination profile and portfolio composition mix.

The model was developed on the in-time sample and validated on the out-of-time sample.
Subsequently as the model was fine-tuned through progressive iterations, few of these variables were either dropped/removed while others were interacted together to capture the model's differentiated sensitivities at granular segment level to key risk drivers, as deemed necessary. Please note that the LGD model does not have input variables called foreclosure timeline or liquidation timeline. However, the model uses state-specific dummy or judicial vs. non-judicial information to account for the loss difference due to state legislation differences around mortgage foreclosures. A comprehensive description has been provided for all finalized model specifications (both PD and LGD). Please refer to Section 5.1.4 for additional details. Additionally, description of modeled outputs (VP, IVP, Delinquency, GCL and NCL) has been provided at the start of Section 6.3 of the MDD.

All the attributes mentioned above are considered critical elements that affect the riskiness of the modeled portfolios. The following rationale provides the additional basis that justifies the consideration of all attributes and connects them to the model's performance

- Loan level attributes (both static and dynamic) are considered important as each loan can behave very differently in different economic scenarios, including voluntary or involuntary prepay and delinquencies.
- Mortgages also have embedded optionality's such as option to prepay (call) or the option to walk away from the loan (put). Since the terms of these options do not generally average out analytically, using loan level specific attributes that analyzes individual loan's behavior provides substantial flexibility in model calibration and specification
- Company specific policy mandates and strategic operational decisions affect the performance of loans. Specific guidelines around loss mitigation policies affect the eventual losses incurred on certain loans
- Bureau attributes provides information on the borrower's overall financial health and is considered useful to assess the potential riskiness of a loan
- Natural disasters can affect the distribution of losses and hence affect the overall riskiness of the loan. Based on model sponsor leadership requests, effects from natural disasters such as hurricanes were included in the list of risk drivers considered in the 2019 model redevelopment project in the form of indicator variables.
- Real estate represents a significant portion of most people's wealth, and this is especially true for many homeowners residing in the United States. According to the US Census Bureau, the number of households in the U.S. has been growing steadily over the past decades, as has the population. The total number of households has doubled from about 63 million in 1970 to more than 126 million in 2017. Based on the latest census report, the homeownership rate, as of the first quarter of 2018 is 64.2%. Given the sheer size and scale of the NA real estate market, it is very important to understand the broader macro-economic factors that affect the real

estate performance, which in turn affects how banks allocate their reserves and capital in support of holding these loans in portfolio. An improving economy is reflective of increasing housing prices, rising consumer confidence (rising stock market indices and lower volatility) and increase in per capita purchasing power which lowers the default rates and corresponding losses. Default, delinquencies, prepayment, and severity appear to be correlated through their joint dependence on common economic factors and hence are considered imperative to the build of robust credit risk models.

Presented below is a short excerpt of the current state of the US Economy-

- Housing Price Index (HPI) - The HPI indicator provides the trend in housing prices over time which heavily influences the changes in the rates of mortgage defaults, prepayments and housing affordability. The HPI witnessed an increase of 9% from quarter four of 2017 to quarter one of 2018.
- Interest Rates - One of the primary ways the Federal Reserve can influence the broader economy is through interest rates. Generally speaking, the lower interest rates are, the easier it is for the economy to grow. But when the economy grows too fast, it can overheat and create the sort of bubbles and subsequent crashes that can undermine long-term growth — like the 2008-2009 financial crisis. In US, the Central Bank aka the Federal Open Market Committee (FOMC), manages the interest rate through the Federal Funds rate. Although the mortgage rate is typically tied to factors like the returns on 10-year Treasury note, the returns on 10-year Treasury note are influenced by the federal funds rate, so the ripple effect of changing interest rates typically makes its way to the mortgage market as well. Given the steady progress in economic growth in the recent past, the FOMC had raised the rates in a staggered way starting from late 2015 till date with the current rate at 1.75% as of the first quarter of 2018. The 10 yr. Treasury yield correspondingly rose to 2.75% in quarter one of 2018 from 2.5% end of Q4 2017.
- Unemployment Rate - Unemployment rate remained steady at 4.1% for the first quarter of 2018. While the drop in unemployment rate might be partially explained by the contraction in the labor force, yet the jobless rate has definitely shrunk to historically low levels. While there have been some uncertainties around the ongoing trade dialogues between the world leaders, economists say it is still too early for employers to initiate any kind of changes to the staff or expansion plans in response to the tariffs.
- GDP - With respect to the GDP rate, this has remained on a steady pace, with the real GDP growing by 2.0%, the first quarter of 2018, fueled by the increase in process of all consumer products, corporate profits and business investment and tax reform initiatives.
- Income – This is the state level aggregated household income which is the total of all incomes in the state without adjustments for inflation, taxation.
- Stock Market Movements [S&P500 & VIX] - The S&P 500 is widely regarded as the best single gauge of large-cap U.S. equities and essentially is the single-most important metric used for evaluating the overall health of the US equity market and associated

market risks. There is over USD 9.9 trillion indexed or benchmarked to the index, with indexed assets comprising approximately USD 3.4 trillion of this totals. The S&P500 index has stayed stable from quarter four of 2017 to quarter one of 2018, reflecting the ongoing strong state of the economy.

Additionally please find below all the macroeconomic attributes that have been considered and eventually used for the 2019 RM model suite. For interest rates, please refer to Chapter 2 of the MDD.

Other macroeconomic factors	Measurement Unit	Geographical granularity	RM
HPI	Average HPI reported monthly	County	HP growth rate and Mark-to-market CLTV in PD and LGD
Unemployment Rate	Average unemployment reported monthly	State	Both level and growth rate used in PD and LGD
GDP	Calculated based on annual GDP and reported quarterly	State	Both level (in only 1 equation) and growth rate in PD
SP500	Average SP500 reported monthly	US	Growth rate in PD
VIX	Average VIX reported monthly	US	Tested but not selected
Income	Calculated based on annual income and reported quarterly	State	Growth rate in PD

- What changes have been implemented from previous versions (if any?)

  [Include the key attributes that differentiate the new model from its previous version (e.g., additional segments were added, granularity increased, groupings were done through different methodology, variables were stationary, hold-out sample length increased). What reasons drove the change, and how does the change impact soundness of the model?]

Presented below is a summarized overview of the targeted model developments for this year's CCAR submission, compared to last year's version. The redevelopment items have been carefully selected based on prior year's models cited limitations, external/internal review & recommendations and CAMU's continued commitment to improve existing processes through self-identified initiatives and maintain highest level of transparency and consistency across the models' reported usages.

- Quantitative segmentation analysis (decision tree) has been used to justify the model's prior year's segmentation scheme and also discover additional new segments of interest. The model's performance has been improved across segments through the addition of interaction attributes and/or implementing a time-decay factor for legacy originations to not penalize their strong performance to-date, including their resilience to the macroeconomic stress (Wall St Originations).
- Comprehensive benchmarking with a structurally different model suite (Hazard vs transition) Proposed alternative Challenger model suite - Model Risk Management is also validating Method C in preparation for the 2019 CCAR submission. Additionally, specific benchmark analysis has been conducted between the 2017 and 2019 model suites to gauge the model's performance(backtest and sensitivity) from the last version. Please refer to Section 6.5 for additional details.
- Aligned product identification logic to be consistent with the proposed business usages of the model.
- Introduced recent period data in the model development pool to factor in the evolving changes in portfolio composition.
- Model usage grid modified to clearly separate out between CCAR and Non-CCAR usages. Documentation would specifically mention that non-CCAR usages (& outputs) are an add-on module over the mandated CCAR requirements to aid business decision.
- Justified model complexity by introducing a systematic approach towards defining modeled vs non-modeled transitions. All transitions go through a waterfall logic wherein they are tested based on their correlation with macro-economic trend, total overall volume, their contribution volume to the source/destination transition cells and modeling soundness (C statistics). All transitions which passed the waterfall were modeled while those not pass the waterfall were non-modeled. The creation of this systematic logic helped in addressing the model's complexity

(and increase in modeled equations) through appropriate justification that leveraged both business intuition and statistical analyses. Transitions, which had scant data, were estimated using the non-modeled transition assumption thus eliminating potential model overfitting.
For non-modeled transitions, a lookup table was created for stress and non-stress, respectively, across some critical segments - In-Trial vs Not –in-Trial Government vs Not-in- Trial - Conventional based on CMI in-sample data. Further the non-modeled transitions were shocked using a 10% rate [based on the Model Testing Guidance] to validate the stability of assumptions to shocks.

- Appropriate justification provided for the coping strategies used for combining the CMI and CPB portfolios together. Justification has been validated using decision tree analysis which did not reflect much dissimilarity across these two portfolios. Coping strategy to combine CPB/CMI, justified by limited data availability. Created CPB specific balance effects and lagged macro-econ variables in the model.
- Addition of new macro-economic attributes to capture the influence of equity market performance on real estate portfolio.
- Introduced new HPI burn-in and VP interest rate burnout attributes to capture the portfolio composition changes in response to the changes in the broader macro-economic environment
- Deep-dive analysis was conducted and variable transformations (change of change, interaction) were introduced to capture the model's stress peak losses and improve the model's VP equation sensitivity to macro-economic factors and to capture the model's recent period BUK7 -> IVP increase.
- Granular level characteristics analysis and reporting were made consistent with the model segmentation analysis for recent and stress periods.
- The model development data spans from 02/2006 to 12/2011 and from 04/2014 to 12/2017, which now include both the historical stress and most recent time periods to factor in one full business cycle and also capture the changes specific to portfolio composition in recent times.
- Replaced the prior DV model using a DV lookup logic that eliminated the prior model's sole dependence on HPI. The haircut logic has been implemented taking into account its differentiated sensitivities to key loan level attributes for all properties with or without an existing distressed valuation. This new logic not only helps remediate the concern with the prior model but also moves towards applying a consistent approach for collateral valuation, as recommended by the business.
- The LGD modeling framework was changed from modeling losses by their disposition types to modeling losses by their outcomes (Full vs partial vs zero) for first lien loans. This change in modeling framework had a three-fold advantage – 1) Zero loses were no longer excluded from the model development sample, especially given the influx in zero losses in recent times due to consistently increasing HPI,2) Removed the uncertainty associated with modeling losses by their disposition types, which are strategic decisions that cannot be predicted in advance, and 3) aligned the modeling approach across first and second liens for a more consistent interpretation and comparison of model results.
- Sensitivity and backtesting have been done at appropriate granular levels as recommended by

model reviewers in compliance with the Model Testing Guidance.

- Production code has been simplified to incorporate the specific changes and enhancements from 2018 IRMO Code Review and removal of redundancies and duplicated code.

5.2.3 Manual Adjustments

[Please provide both technical and business rationale with analytical evidence in support of the manual adjustment in any part of the model processing components. A 'manual adjustment' is any manual change made to the model algorithm or model estimates that is transparent, logical and repeatable.]

- If manual adjustments have been applied, are they algorithmic, transparent and repeatable?

Not Applicable. No manual adjustments have been applied. Any proposed manual adjustments to the model post MDD submission would be discussed thoroughly with the Model Sponsor and End Users and proposed in the Overlay documentation, which is also independently validated.

- Are the adjustments in the estimated parameters of the model justified?

Not Applicable. No manual adjustments have been applied.

RM Model Eligibility Assessment (MEA) Responses

1. Model Sponsor is required to provide the rationale for modelling these transitions. For a few PD and LGD, i.e. Second Lien-Stage 1 - Full Vs Partial, it is observed that 1). the intercept is the most significant factor to the model, which could mask the effect of true risk drivers, 2). the coefficient of intercept is highest among all the variables, which makes the model not be able to differentiate the risk between individual loans. Model Sponsor is required to justify this.

   Answer - The test of the intercept in the procedure output tests whether this parameter is equal to zero. For a logistic model, zero intercept implies that the event probability is 0.5. This is a very strong assumption and is most likely not the case in PD and LGD stage 1 models. So, a highly significant intercept in some of these models should not be a problem. The characteristic analysis also shows reasonable risk ranking among major risk drivers.

2. Model Sponsor is required to provide the rationale for modelling these transitions. High R-square is observed for LGD Stage 2 model, 94%, which may indicate potential over-fitting.

   Answer - High R-square not necessary indicates over-fitting. The stability test (table 6.2.2.1) on LGD stage 2 model shows that the model is robust and stable in both In-time out of sample (OOS) and out-of-time sample (long). Therefore, the model is not over-fitting.

5.3 Model Assumptions

This section will provide additional information to the "Automated Document" Section A.3.: "Inputs, Outputs & Assumptions" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

[Provide description and testing results of all model assumptions, including assumptions related to the chosen modeling technique. State whether the estimation procedure used required any specific mathematical assumptions, and how such assumptions were satisfied. For example, for competing risk models (i.e. default vs. prepay/closure), if Begg-Gray normalization of sequentially-estimated multinomial probabilities was used in lieu of simultaneous estimation of those probabilities, explain how the assumption of independence of irrelevant alternatives is not violated.

List and describe all explicit and implicit assumptions required for the model to be sound. Describe the governing statistical and structural assumptions. For each modeling assumption listed, refer to Section 6 on Model Testing to provide description of how the assumption was evaluated. If assumption is not "testable", please explicitly state as such.]

All model assumptions are considered well justified in the context of the modeling objective and the available resources (data, tools, system, software, human capital, timeline, etc.) at CAMU's disposal. These assumptions are also consistent given the portfolio characteristics and the strategic oversight maintained for these portfolios. The sensitivity of the model results to changes in key assumptions, which includes shocking the non-modeled transition rates by 10%, has been discussed in MDD Chapter 6.

- The model is dependent on future macro-economic values. Even if it were perfect, the accuracy of the model is only as good as that of macro-economic forecasts used to operationalize it. The sensitivity to macro-economic factors/assumptions has been tested in Section 6.4
- As reiterated within Sections 3.1 and 3.2, CAMU took a systematic approach based on empirically derived waterfall logic towards estimating the non-modeled transitions rates. The non-modeled transition rate assumptions are based off on the historical data, split across stress and non-stress periods and some other key segments which have sufficient differentiation of sensitivities to either risk factors/macro-economic indicators. For additional details on the Waterfall logic and the preferred approach utilized for non-modeled transitions, please refer to Section 3.1 and 3.2 of the MDD. Further, based on the Model Testing Guidance Requirements, all non-modeled assumptions have been tested using a 10% shock rate. The sensitivity of the non- modeled transition assumptions has been tested in MDD Section 6.4.

5.4 Model Limitations

This section will provide additional information to the "Automated Document" Section A.4.: "Limitations" (refer to the "Note on Model Risk Management – Document Automation", in the beginning of this document)

[All models have some degree of uncertainty and inaccuracy. These can sometimes be quantified. At other times, only a qualitative assessment of model uncertainty and inaccuracy is possible.]

What are the key limitations of the approach used?

Model uncertainty is a commonly encountered ubiquitous effect in the realm of econometric modeling. While the robustness or integrity of the model should remain largely unaffected given the uncertainties, CAMU acknowledges the fact that it is necessary to methodically enumerate the different sources of model uncertainties or limitations in a coherent structure to ensure transparency across all modes of communication

The following list summarizes the list of model limitations for the 2019 Method A Residential Mortgage Model:

- **Dependency on future macroeconomic values:** The accuracy of the model is dependent upon the macroeconomic forecasts used to operationalize them. Since macroeconomic assumptions are important inputs to the model, the model output accuracy is critically dependent upon future macroeconomic assumptions made and provided by The Global Country Risk Management Office (GCRM).

  Mitigation: CAMU has utilized all macro-economic files as was provided by the GCRM team. To ensure integrity of the data used, appropriate testing and quality checks were done to ensure the quality and materiality of the data met the requirements for CCAR usages. Additionally to ensure the same set of data was used for model development and production, all required input files were copied to a frozen dataset.

- **New origination loans:** The model is estimated and implemented as a transition model. It doesn't have a mechanism to predict the volume of new origination in the forecasting window. So it can only predict the future loss of the portfolio as of the snapshot date. Any loss associated with new bookings cannot be realized in the model.

  Mitigation: Mortgages are long lived assets and the vast majority of losses come from seasoned loans, not new originations. Furthermore, end-users of the model have the ability to make adjustments/overlays to account for new loan losses.

- **Expected Probabilistic Approach -** There are generally two approaches to implement the transition rate models. One is to use Monte Carlo simulation in which multiple paths are run. Each path follows a single transition based on probability weighting. Under different runs, different transitions may be taken. This method requires the numbers of runs to be sufficient enough to achieve convergence. An alternative simplified approach of calculating aggregate balances by status vector can be applied when the models have a transition feature (i.e., the forecasted probabilities in next month depend on the loan status in current month but not on the entire history of loan statuses). That is, in each month, all the loan balances in a specific bucket can be added for future forecasting.

***Mitigation*:** The model herein employs the second approach for the Model implementation. The status vector approach maintains the number of loan statuses constant over time. In each month, the expected balances rolling into each bucket per account are aggregated and used as the basis for future forecasting. Further the model estimates remain unbiased and some payment history variables are used to make the model more accurate despite potential limitations in their simulation method.

- **Impact of loan modifications or loan sales** - The performance of the loans after modification or sale is either difficult to predict or even impossible to track since they are policy driven initiatives. CAMU has experimented with several different types of approaches to simulate the action loan It is important to note, it is always assumed that there were no future modification or assets sales in the forecasting.

  Mitigation: For the 2019 backtesting, CAMU, as recommended by MRM, would use the naïve censoring approach for backtesting of the action loans. For additional details on the censoring approaches, please refer to Section 1.5 of the MDD. Further the model has the built in features and flexibility to add adjustments based on expected future modification activities and asset sale strategies.

- **Ancillary Costs associated with Loan Foreclosure** - The target variate for the LGD model reflects only the loss given default (LGD) for loan loss calculation, i.e., NCL or ENR loss. It does not include any costs or expenses incurred by collections, foreclosure, collateral disposition, lost interest, etc.

  Mitigation: Collection costs or other ancillary costs associated with the management with a defaulted loan are dictated by Bank's default management policies and the State's regulation around the same. As such, these policies dictated costs and expenses tend to vary across systems, processes and regions and do not impact the actual losses themselves. Hence they are kept outside the scope of the loss estimation.

It is important to note that this proposed model attempted to address all prior model limitations noted by independent validators – which is discussed in detail in MDD Chapter 2 above.

What are the potential weaknesses in model performance?

Mathematical models, directly or implicitly, are limited in the sense that they must be tractable, which is to say, they are useless unless they can produce insightful results. The "math of reality", is remarkably complex, and can only be faithfully duplicated in an abstract model of human design, if all underlying processes and data work in a logical way. However, if any of the inherent logic or data trends are substantially different from the reality they purport to represent, the model results will not be insightful.

The following list represents the model's perceived weaknesses, as self-identified by CAMU for the 2019 CCAR process. Please note a full discussion and analysis of the model's performance is illustrated in Chapter 6 of this MDD.

- **IVP Prediction Error** - An involuntary payoff happens as the last resort for borrowers who are unable to meet their monthly mortgage obligation and the lender initiates the proceedings to charge off the account. Once the loan hits the IVP status, the next steps of action and the loan's subsequent status relies less on the borrower's credit quality or loan performance and more on the given company's foreclosure and loss mitigation policies and the state's (loan's state of origination) regulation around the foreclosure process. Typically states which have a court-adjudicated foreclosure process experience an increasing level of costs, given the length of the foreclosure timeline. Apart from the state-mandated foreclosure laws, enactment of moratoria acts by the major financial institutions in US, including BankGroup, [https://www.wsj.com/articles/SB123454524404184109,https://www.Bankgroup.com/Bank/news/2009/090213a.htm,https://blogs.wsj.com/washwire/2009/02/13/chase-Bank-to-implement-foreclosure-moratoriums/?mod=article_inline] was the additional commitment provided by the financial institutions to work with the government towards creating a financial stability plan aimed at keeping people in their homes. As part of the moratorium memorandum, borrowers were granted permission to reside in the defaulted property for a specific period of time, without the lender initiating any foreclosure processing. The moratorium is a grace period of sorts during which the borrower gets respite from the monthly payments, even though interests tend to accrue during the moratorium period. The government still continues, to this day; administer the moratorium acts, as a meaningful way of helping home owners in distress [https://www.hud.gov/press/press_releases_media_advisories/HUD_No_18_039]. In addition, to the moratoria acts, enactment of the anti-deficiency judgment laws [https://realestate.findlaw.com/foreclosure/what-are-anti-deficiency-laws.html], post the Financial Crisis of 2008-2009, is the most modern demonstration of the government's desire and intervention to protect the mortgage borrower.

  Given the on sleuth of policy intervention (internal + external) around the defaulted loans, it has been a relentless modeling challenge to model these IVP transitions accurately. Since policies are more driven by strategic initiatives (either at the Federal/State level or the specific company level), it is not possible to capture the effect of policy changes or the drivers behind these policy changes using a statistical model. To provide some context, loan's that were relegated to the Moratorium status at the time of the Crisis (2009) had no IVP, which were in complete contradiction to the prevailing macro-economic conditions within the broader economy (very high unemployment, low income/GDP, etc.). Moreover, such policies often trigger moral hazard because they encourage even solvent homeowners to seek debt forgiveness (Mayer et al. (2014)). Given this idiosyncratic behavior associated with the IVP loans, there was no meaningful approach to accurately model these loans. Although considerable work is being carried out within the realm of probability modeling in forecasting policy effects that can address the breakdown of equations and/or unreasonable forecasts; it is not yet standard for macroeconomic theorists to fully explore the behavior of their models in response to policy intervention. As, such, CAMU would like to disclose this as a model weakness and would seek to collaborate with model reviewers, sponsor and business users on the best possible ways to model these loans.

- **Increase in Government BUK7(180+DPD) -IVP transition rate in 2016/2017** - A government-backed loan is a loan subsidized by the government, which protects lenders against defaults on payments, thus making it a lot easier for lenders to offer potential borrowers lower interest rates. Its primary aim is to make home ownership affordable to lower income households and first-time buyers. While there are many different types of government loan programs, the two commonly funded ones pertain to FHA (backed by the Federal Housing Administration (FHA), a government agency) and VA (backed by the Department of Veteran Affairs, which allows zero down financing). The Government loans funded by BankGroup are typically characterized as having 97 LTV at origination with corresponding high DTI (Debt-to-income) ratio. Policy wise, when a government loan becomes IVP and goes into subsequent foreclosure sale; and if BankMortgage wins the bid during the foreclosure induced sale; this loan would tend to remain active on the Bank's system (data warehouse) until conveyed to HUD (Dept. of Housing & Urban Development). But the loan even though showing 'active' status is not a real loan as the property was already foreclosed and the collateral is owned by BankMortgage. In particular, the 180+DPD to IVP rate for government loans has been increasing in recent time periods - 2016 and 2017 which contradicts with the prevailing macro-economic trends which are marked with significant and consistent home price appreciation and unprecedentedly low unemployment rates. Further, during the same timeframe, BankMortgage initiated a transfer of middle office functionalities, who managed the foreclosure of the government loans to a different site location. This transfer of personnel led to instability in the management process and the subsequent issuance of MRA's from OCC around the management of debts previously contracted (DPC). Although CMI had increased performance since then, this also played a role in the uptick of the IVP rates in the recent timeframes. Again given the outcome incommensurability (change of outcome across transitions are not possible to capture without additional conjecture around Bank's existing processes and policies around them), it is not possible to accurately model the government loan IVP performances for recent periods. Again, to ensure maximum transparency within the modeling process, CAMU would like to call out this model weakness as part of the 2019 model submission.

- **Loan Curtailment -** Loan curtailment shorten the life of the loan by paying off the balance ahead of schedule. Principal curtailment is when the borrower makes extra principal payments in an effort to reduce the outstanding principal balance. This can also be called a partial curtailment. A total mortgage curtailment is when the entire balance of the loan is paid off with a large lump sum ahead of schedule1. A deep-dive analysis and review of the 2017 Method A production code by the model reviewers' revealed that amortization calculations within the Method A model suite assume no curtailment logic and instead are just based on multiplying a factor by the prior months balance which doesn't necessarily equate to borrowers' actual payment when a curtailment happens.

  Within the 2019 production package, CAMU has improved the amortization module that will utilize the contractual payment of the loan and the starting balance including curtailment, as of the portfolio snapshot date. For additional details, please refer to attachment –' 1.1 CAMU Response Draft_NA Mortgage Model Execution Code_GOLD COPY'.xlsx. As such, curtailment is considered in the model as of the snapshot date. However, please note the model is not capable of forecasting future curtailments, since this lies outside the scope of Model Usage Grid.

What are the internal and external factors used to evaluate model weaknesses?

The list of internal and external factors that were considered when evaluating model weaknesses/ limitations include changes in the model usage, changes in the portfolio composition, changes in credit policy and business strategy, industry and economic environment changes, regulatory environment changes, new MRAs, model validation findings, and observed data errors or other operational failures.

The overall model evaluation is based on three criteria: accuracy of backtest results, sensitivity results, and characteristic analysis results. The accuracy of backtest results includes end to end loss accuracy, accuracy of loan counts for each bucket (especially defaults), and if the loan counts and balance predictions align with each other over time. Sensitivity looks at the responsiveness of model output to the macro-economic drivers. Characteristic analysis is backtest accuracy at the attribute or other key dimensional level, as deemed noteworthy via the model segmentation scheme.

Are there any known limitations on the range of the model outputs outside of which the model might not perform as expected?

The model was developed using a wide range of performance data both temporally and cross-sectionally. The logistic structure of the model ensures that model incidence outputs don't fall outside the range of probabilities. As such there are no known limitations on the range of model outputs. While the model results has been primarily evaluated over short (12 months), medium (27 months) and long (60 months) term performance periods, it can produce results for the lifetime of the loan. However, as the horizon extends there is an expected decrease in forecast accuracy, this can even be seen by the testing guidance policy thresholds loosen with added months in the horizon.

Are any potentially material risks missing including those in the Material Risk Inventory?

Example: *For deals beyond 10y, stochastic interest rates may become materially significant.*

No. The robust variable selection and generalized nature of the modeling framework suggests no material risks are missing. CAMU works closely with other model sponsor teams (including model users and stakeholders) to review Material Risk Inventory items that can be applicable to models managed by the team.

What factors, if any, mitigate the limitations of the output?

[Include a sub-section for each identified limitation of the output (e.g., historical bias, failure to pass statistical tests etc.) and how those limitations are mitigated. If none exists, provide a rationale for why one is not needed.]

Example: *The business was engaged in the variable selection process and has selected variables where they believe that historical relationships will continue. This is intended to mitigate historical bias.*

All potential model limitations have been outlined above with corresponding mitigation plans. Please note the redevelopment of this model attempted to address all prior known limitations noted by CAMU as well as independent reviewers and validators.

6. Model Testing

[Model testing includes checking the model's accuracy, demonstrating that the model is robust and stable, assessing potential limitations, and evaluating the model's behavior over a range of input values. It should also assess the impact of assumptions and identify situations where the model performs poorly or becomes unreliable].

The purpose of this chapter is to evaluate whether the model performs as intended. For this purpose, a range of testing exercises should be performed on the subject model and documented along with the interpretation of the test results. The preferred tests should be based on the model methodology and should be commensurate with the applicable principles as per the Model Testing Guidance. All assumptions made during model estimation (listed in Section 5.3) should be thoroughly tested and documented in this Section. Please refer to the Model Testing Guidance for the list of mandatory tests to be performed as per the model methodology. Also, the Model Sponsor is required to provide all the testing results in line with the model usage and as per the model testing guidance. In case any of the mandatory tests are not performed, appropriate rationale must be provided by the Model Sponsor/ Developer and additional tests, if deemed appropriate for a particular model or scenario can be performed.

Please reference the challenges identified in Section 3.2 and please provide the alternative quantitative analysis and/or qualitative justification conducted in lieu of the performance test.

For loss projection purposes, testing should be performed on individual default/loss components through to GCL and NCL level and as per the model usage. However, if the MDDT states that the model presented in MDDT only covers a specific component of the overall NCL (or GCL) Forecasting Model, the entity that is responsible for ensuring the full suite of testing and documentation through NCL (or GCL) must be attached in a separate document to demonstrate complete NCL Forecasting Model testing and documentation. In the case of nested models, each component model should be individually tested.

The performance testing of individual models can be documented in one of two ways:

- If the individual model has a standalone MDDT, please reference the name of the MDDT, reference the specific MDDT section number, and provide a brief summary of performance test results
- If the individual model is included as part of this MDDT and does not have a standalone MDDT, please conduct the tests outlined below (as applicable) and describe results

[Please note that this section is applicable to internally built as well as vendor models.]

6.1 Diagnostic and Statistical Tests

[Describe the model specific statistical tests and binding requirements that were performed to ascertain the soundness of the model.

In addition, any other tests performed to assess the Soundness of the Modeling Approach should be documented in the last paragraph of this section (Section 6.1.9 Other tests – Model-specific diagnostics. Potential limitations for any additional assumptions have to be listed in case these were made. Please refer to the Model Testing Guidance for testing requirements.]

The North America CCAR Loss Forecasting Models is comprised of multiple models and corresponding logics used to support the end to end framework proposed. While some are statistical models, others are empirically derived lookup logics.

The following represent the complete model suite of the 2019 Method A residential mortgage package-

1. PD Transition Model: Depicts the transition flows derived both from statistical models and empirically observed historical estimates

2. LGD Two Stage Model: Two stage model that estimates the losses by their respective outcome and then models the probability of partial losses

To evaluate whether models perform as intended various statistical tests and validations are performed on the model inputs and outputs-

- Independent variables have been tested for their significance based on the p-values, their alignment with the perceived business intuition and multicollinearity using the Variance Inflation Factors (VIF)
- The predictive accuracy of the model has been tested using the predicted cumulative error, RMSE, COV and AIC statistics
- The model was also tested for its robustness to ensure parameter stability across short and long samples through extensive out of sample and out of time related tests
- Finally to ensure that the model remained sensitive to macroeconomic scenarios, sensitivity tests were performed on stressed macroeconomic conditions for base and stress (severely adverse, as defined by FRB) scenarios

All model testing results comply with the MRM's CCAR Model Performance Testing Guidance and are considered to be sufficient enough to establish the model's accuracy and demonstrating the soundness of the model for both CCAR and non CCAR initiatives. The Diagnostic (Section **Error! Reference source not found.**) and Model Performance (Section **Error! Reference source not found.**) sections will focus on the models diagnostics. Back testing (Section **Error! Reference source not found.**) and sensitivity testing (Section 6.4) will focus on the entire model package's end-to-end results, attesting to the model's performance.

6.1.1 Summary of diagnostic and statistical test results

[Provide summary results of key tests performed in this section. State if the test results are acceptable or not with respect to the model type/ model components and why.

Explain the results and describe subsequent limitations to the model usage and mitigation actions for model use in production in the table below.]

The Method A Residential Mortgage PD model (28 individual PD model equations) satisfied all statistical tests conducted according to MRM model performance testing guidance, and are proven to be statistically sound.

The modelling team performed the following diagnostic tests on each single PD equation

1) Statistical Significance: The model development team elected to use the Wald Chi-Square test to determine whether the parameters are statistically significant. All the independent variables considered within the model are highly statistically significant with P-value<0.05, in fact most of them had P-value < 0.0001 on both development and validation samples.
2) Business intuition: The statistical sign and estimate was compared with the observed business process to align the parameter's statistical significance to its business intuition.

3) Multicollinearity: The model development team used the Variance Inflation Factor (VIF) as the testing criteria to detect the presence and degree of multicollinearity. The criteria that VIF should be less than 5 was adhered to during the model development process. Please notice that it is reasonable and acceptable if observing high VIF between spline or interaction variables with their raw variable.
4) Parameter analysis among development and validation samples: Method A modeling team compared the parameters estimated on model DEV data, INT data and a full sample that includes DEV, INT and OOT sample. Variables are kept only when their estimates are significant (P-value <=0.05) and have consistent sign and meanwhile do not vary largely among different samples. More details will be discussed in section 6.2 Model Robustness and Stability.

Table 6.1.1.1 below summarizes all test types and test results. Details will be discussed in section 6.1.2 to 6.1.9. For the detailed statistical tests results including parameter significance, parameter stability analysis and VIF for the RM PD model, please refer to the attachment excel files 6.1 Residential Mortgage Model Diagnostic and Statistical Tests.xlsx. For the detailed statistical tests results including parameter significance, parameter stability analysis and VIF for the Severity model, please refer to the attachment excel files –'6.1 Severity Model Diagnostic and Statistical Tests'. The detailed justification for all parameter signs can be referred in attachment excels- 'Residential Mortgage PD Spec Justification.xlsx' for the PD model. The LGD model specifications can be found in the attachment - '5.1.4 Severity_model_specifications.xlsx'.

Table 6.1.1.1 Diagnostic and Statistical Test Results for Residential Model

Model component/ segment	Type of test	Test results	Business justification/ Limitations	Mitigation
PD model	1. Parameter significance 2. Parameter sign economic meaning 3.Multicollinearity 4. Development vs. in-time or full sample parameter analysis	Satisfactory	No major limitation	N/A
LGD model	1. Parameter significance 2. Parameter sign economic meaning 3.Multicollinearity 4. Development vs. in-time or full sample parameter	Satisfactory	No major limitation	N/A

	analysis			

Note: EAD and DV and other related assumptions / logic are discussed in detail in MDD sections 5.1 and 5.3 above.

6.1.2 Statistical Significance of Parameter Estimates

To the extent applicable, provide evidence that the model coefficients corresponding to selected independent variables are statistically different from zero

[Provide the results of the coefficient significance tests. Comment on the results. If not tested, please provide appropriate rationale.]

A Wald Chi-Square test is used to test the statistical significance of each coefficient (β) for PD models. The test is calculated as

$$W = [\beta / SE_{\beta}]^2$$

Chi-Square and Pr > ChiSq are the test statistics and p-values, respectively, testing the null hypothesis that an individual predictor's regression coefficient is zero, given the other predictor variables are in the model. The Chi-Square value follows a central Chi-Square distribution with degrees of freedom given by DF, which is used to test against the alternative hypothesis that the estimate is not equal to zero. The higher Wald Chi-Square value of a variable, the better it is.

PD Model Parameter Estimates

The Method A modeling team entirely re-estimated the 2019 Method A Residential Mortgage PD model, which includes 28 equations. All the independent variables selected in each individual model are statistically significant with p<0.05, most of them in fact have a p-value <0.0001. The strictly small p-values contributed to building robust models and minimized over fitting.

The table 6.1.2.1 illustrated below are the statistical significance tests for the parameter estimates for equation bucket01 (0~29DPD) to bucket 2 (30~59DPD) estimated by logistic regression. It serves as an example of evidence for the statistical significance of parameter estimates. For the complete details of parameter estimates statistical significance of all PD models, please refer to the **P-value** tab of the attachment excel files “6.1 Residential Mortgage Model Diagnostic and Statistical Tests.xlsx.”

LGD Model Parameter Estimates

Tables 6.1.2.2 to 6.1.2.5, has the LGD stage 1 and stage 2 models estimates for first lien loans respectively. Please refer to the tab – ‘6.1.2 Coefficients’ of the attachment excel files –‘6.1 Severity Model Diagnostic and Statistical Tests’ for the LGD second lien stage 1 and stage 2 model diagnostics results.

Table 6.1.2.1 Parameter Statistical Significance for RM Buk01->BUK2 Model

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	2.954	0.107	764	<.0001
B_S_BROK_IND	0.322	0.012	766	<.0001
B_S_CORR_IND	0.554	0.003	27193	<.0001
B_S_WLST_IND	0.690	0.005	20967	<.0001
B_S_LoanPurpPurch_IND	0.049	0.003	251	<.0001
B_S_LoanPurpWork_IND	0.922	0.026	1298	<.0001
B_S_OCC_IV_IND	0.315	0.007	1939	<.0001
B_S_lowIncDoc_IND	0.303	0.005	3965	<.0001
B_S_pmi_IND	0.159	0.007	495	<.0001
B_S_DocMISS_IND	0.158	0.003	2621	<.0001
B_S_LoanPurpMiss_IND	0.401	0.019	437	<.0001
B_M_coborrower_IND	-0.245	0.002	11063	<.0001
ARM_5Yr	-0.267	0.007	1676	<.0001
ARM_GT5Yr	-0.318	0.009	1405	<.0001
P_S_OrigIntSpread_FRM	0.081	0.001	10991	<.0001
B_S_INT_ONLY_IND	-1.215	0.037	1090	<.0001
POST_IO_IND	0.148	0.015	97	<.0001
P_AND_I_ratio_IO	0.600	0.031	367	<.0001
POST_ARM_6M_IND	0.198	0.012	265	<.0001
POST_ARM_IND	-0.081	0.006	199	<.0001
D_M_PRIN_BAL_LE2K	0.969	0.011	7616	<.0001
Balloon_Mature	1.458	0.036	1600	<.0001
Katrina	0.098	0.014	50	<.0001
Sandy	0.212	0.032	43	<.0001
FEB	-0.248	0.004	3149	<.0001
MAR	-0.098	0.004	558	<.0001
S_M_FicoRefresh	-0.010	0.000	234219	<.0001
S_M_FicoRefresh_SP640	-0.010	0.000	13562	<.0001
S_M_FicoRefresh_SP720	0.009	0.000	2487	<.0001
B_S_missS_M_FicoRefresh	0.240	0.007	1328	<.0001
N_M_age_oldest_trd	-0.001	0.000	2065	<.0001
miss_N_M_age_oldest_trd	0.065	0.003	525	<.0001
Post2010_Orig_CONV	-0.566	0.013	2023	<.0001
HPI_tm12_ratio_LSP1_CMI_CONV	-0.552	0.010	2941	<.0001
MTM_CLTV_SP40_1st_CONV	0.004	0.000	1691	<.0001
MTM_CLTV_SP80_1st_CONV	0.003	0.000	717	<.0001
MTM_CLTV_SP80_2nd_CONV	0.005	0.000	5571	<.0001
post_neg_CONV	-0.214	0.009	609	<.0001
R_M_State_UnempB12M_IO_CONV	0.655	0.014	2088	<.0001
R_M_State_UnempB12M_BROK_CONV	0.107	0.010	114	<.0001
income_12m_ratio	-0.576	0.041	200	<.0001
N_M_MOB_cap_WLST_CONV	-0.002	0.000	972	<.0001
P_M_CURR_NOTE_RATE_ARM_CONV	0.071	0.001	12926	<.0001
D_M_PRIN_BAL_in10K_CMI_CONV	0.002	0.000	467	<.0001
P_M_State_Unemp_Rate_lag12_GOV	0.026	0.002	193	<.0001
MTM_CLTV_SP80_GOVGT3	0.009	0.000	3681	<.0001

Table 6.1.2.2: First Lien Model - Stage 1 – Zero vs Partial - Regression Coefficient Estimates

First Lien-Stage 1: Zero Vs Partial							
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)	Variance Inflation
Intercept	1	-0.2522	0.1141	4.8898	0.027	0.777	0

LTV_mtm_SP_gt20_le100	1	-0.0383	0.00136	790.8103	<.0001	0.962	1.08915
log_HPI_tm12_ratio	1	3.6738	0.317	134.286	<.0001	39.401	1.08915

Table 6.1.2.3: First Lien Model - Stage 1 – Full vs Partial - Regression Coefficient Estimates

First Lien-Stage 1: Full Vs Partial							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**	**Exp(Est)**	**Variance Inflation**
Intercept	1	57.3267	1.8698	939.9726	<.0001	7.88E+24	0
log_curr_bal_SP_le300k	1	-1.6178	0.0249	4224.9089	<.0001	0.198	1.4448
log_curr_bal_SP_gt300k	1	-3.5151	0.1537	523.2041	<.0001	0.03	1.34551
log_HPI_tm12_ratio_SP_le0	1	-1.6108	0.1591	102.4582	<.0001	0.2	1.1327
UnempRate_SP_gt9	1	0.3786	0.0104	1316.3381	<.0001	1.46	1.16577
Judicial	1	0.4412	0.0297	220.2912	<.0001	1.555	1.15709
HLC_ind	1	1.5806	0.0824	367.5526	<.0001	4.858	1.01008
FHL_ind	1	0.923	0.0877	110.8581	<.0001	2.517	1.00854

Table 6.1.2.4: First Lien Model Stage 2 – Partial Loss Rate - Regression Coefficient Estimates

First Lien-Stage 2: Partial loss rate								
Parameter	**Estimate**	**StdErr**	**Alpha**	**LowerCL**	**UpperCL**	**tValue**	**Probt**	**Variance Inflation**
b_Intercept	2.7847	0.0908	0.05	2.6067	2.9626	30.67	<.0001	0
b_log_curr_bal	-0.401	0.00647	0.05	-0.4136	-0.3883	-61.97	<.0001	1.36528
b_LTV_MTM_SP_gt90	0.0124	0.000296	0.05	0.0119	0.013	42.05	<.0001	3.46141
b_pmi	0.5388	0.0218	0.05	0.496	0.5816	24.69	<.0001	1.01878
b_judicial	0.1877	0.00894	0.05	0.1702	0.2052	21	<.0001	1.09392
b_BROK_IND	0.1433	0.0107	0.05	0.1222	0.1643	13.35	<.0001	1.06881
b_prin_reduction	0.4032	0.1115	0.05	0.1846	0.6217	3.62	0.0003	1.00105
b_UnempRate	0.0351	0.00197	0.05	0.0313	0.039	17.85	<.000	4.29802

First Lien-Stage 2: Partial loss rate								
Parameter	**Estimate**	**StdErr**	**Alpha**	**LowerCL**	**UpperCL**	**tValue**	**Probt**	**Variance Inflation**
							1	
b_log_HPI_t_orig_ratio	-0.2892	0.0331	0.05	-0.3541	-0.2244	-8.74	<.0001	1.44561

6.1.3 Multi-collinearity

To the extent applicable, provide evidence of absence of multi-collinearity.

[List the tests of multi-collinearity conducted. For the test requirements refer to the Model Testing Guidance. Comment on the results of the tests. If not tested, please explain.]

Multicollinearity means two or more explanatory variables in one regression are highly correlated. Multicollinearity is problematic, because it can increase the variance of the regression coefficients, making them unstable and difficult to interpret, for example the variables that are supposed to be significant to become insignificant or the sign of variable to change.

Variance inflation factor (VIF) is an index used to quantify the severity of multicollinearity. It measures how much the variance of estimated parameters increases because of collinearity. Higher VIF indicates higher multicollinearity. CAMU uses a VIF<5 as general criteria within all models to comply with MRM testing guidance.

PD Model Test of Collinearity

All of 2019 Method A Residential Mortgage PD models have max VIF < 5, passing the multicollinearity test. Note exceptions are made if the transformed variable, such as splines or interaction variables, and their raw variables co-exist in the same model. Transformed variables was used in the model to capture the sensitivity of a specific segment or within a specific range of the raw independent variables.

The table 6.1.3.1 illustrated below are the VIF for the parameters in equation bucket01 (0~29DPD) to bucket 2 (30~59DPD) as an example for the RM PD model. As shown below, high VIFs is observed on interested only loan indicator (B_S_INT_ONLY_IND) and the payment shock magnitude of IO (P_AND_I_Ratio_IO). This is because these two variables are correlated with each other as for all NON-IO loans (when B_S_INT_ONLY_IND=0), P_AND_I_RATIO_IO is always 0. But these two variables are both kept in the model to capture the effect that during the interest only period, because of lower monthly payment, IO loans is usually less risky compared to the NON-IO loans. But the delinquency rate with IO loans will increase when IO reset happens and the higher the magnitude of the payment shock, the riskier it will be.

BUK01->BUK2 model also uses refresh FICO (S_M_FicoRefresh) and its splines in the model, in order to capture the non-linear relationship between delinquent risk and FICO. As the FICO splines were created as part of the raw FICO variables, they are highly correlated with each other.

In below table, P_M_State_Unemp_Rate_lag12_GOV and MTM_CLTV_SP80_GOVGT3 VIF are 5.9 and 5.1 respectively. This is because the dataset is dominated by conventional loans, and on conventional loans, the above two attributes, which were created to capture GOV specific effect, all equal to zero. Correlation of these two fields, if limited to only GOV loans, is only 0.2873.

The detailed VIF for individual PD models can be found in Multicollinearity tab of attachment excel files '6.1 Residential Mortgage Model Diagnostic and Statistical Tests.xlsx'.

6.1.3.1 Multicollinearity diagnostic (VIF) for BUK01->BUK2

Variable	Variance Inflation
Intercept	0.0
B_S_BROK_IND	16.1
B_S_CORR_IND	1.5
B_S_WLST_IND	3.1
B_S_LoanPurpPurch_IND	1.4
B_S_LoanPurpWork_IND	1.0
B_S_OCC_IV_IND	1.0
B_S_lowIncDoc_IND	1.2
B_S_pmi_IND	1.1
B_S_DocMISS_IND	1.5
B_S_LoanPurpMiss_IND	1.0
B_M_coborrower_IND	1.0
ARM_5Yr	2.1
ARM_GT5Yr	1.8
P_S_OrigIntSpread_FRM	2.8
B_S_INT_ONLY_IND	83.8
POST_IO_IND	1.3
P_AND_I_ratio_IO	65.3
POST_ARM_6M_IND	1.1
POST_ARM_IND	1.9
D_M_PRIN_BAL_LE2K	1.0
Balloon_Mature	1.0
Katrina	1.0
Sandy	1.0
FEB	1.0
MAR	1.0
S_M_FicoRefresh	8.1
S_M_FicoRefresh_SP640	17.9
S_M_FicoRefresh_SP720	6.5
B_S_missS_M_FicoRefresh	1.1
N_M_age_oldest_trd	1.2
miss_N_M_age_oldest_trd	1.6
Post2010_Orig_CONV	1.6
HPI_tm12_ratio_LSP1_CMI_CONV	2.1
MTM_CLTV_SP40_1st_CONV	5.3
MTM_CLTV_SP80_1st_CONV	12.5
MTM_CLTV_SP80_2nd_CONV	5.2
post_neg_CONV	1.1
R_M_State_UnempB12M_IO_CONV	18.8
R_M_State_UnempB12M_BROK_CONV	16.0
income_12m_ratio	1.3
N_M_MOB_cap_WLST_CONV	2.6
P_M_CURR_NOTE_RATE_ARM_CONV	3.5
D_M_PRIN_BAL_in10K_CMI_CONV	1.8
P_M_State_Unemp_Rate_lag12_GOV	5.9
MTM_CLTV_SP80_GOVGT3	5.1

LGD Model Test of Collinearity

All of 2019 Method A LGD models have max VIF < 5, passing the multicollinearity test. The Table 6.1.3.2 – 6.1.3.3 illustrated below presents the VIF for the Stage 1 and stage 2 models for all liens. The detailed VIF for LGD models can be found in '6.1.3 VIF' tab of attachment excel file –'6.1 Severity Model Diagnostic and Statistical Tests'.

Table 6.1.3.2: Variance Inflation Factor: First Lien and Second Lien - Stage 1 models

First Lien-Stage 1: Zero Vs Partial	
Variable	Variance Inflation

Second Lien-Stage 1: Zero Vs Partial	
Variable	Variance Inflation

Intercept	0
LTV_mtm_SP_gt20_le100	1.08915
log_HPI_tm12_ratio	1.08915

Intercept	0
cLTV_MTM_unemp_down	1.00392
REL_ind	1.05922
HLC_ind	1.05946

First Lien-Stage 1: Full Vs Partial	
Variable	**Variance Inflation**
Intercept	0
log_curr_bal_SP_le300k	1.4448
log_curr_bal_SP_gt300k	1.34551
log_HPI_tm12_ratio_SP_le0	1.1327
UnempRate_SP_gt9	1.16577
Judicial	1.15709
HLC_ind	1.01008
FHL_ind	1.00854

Second Lien-Stage 1: Full Vs Partial	
Variable	**Variance Inflation**
Intercept	0
log_curr_bal_SP_gt10k_le150k	1.48705
log_curr_bal_SP_gt150k	1.28308
junior_ratio_SP_lep20	1.17803
log_HPI_tm12_ratio	1.58909
R_M_UnempB12M_SP_gt1_le1p2	1.59065
lpi	1.04207
judicial	1.02298
wasUSRB	1.03513
REL_ind	1.07291
prin_reduction	1.00404

Table 6.1.3.3: Variance Inflation Factor: First Lien and Second Lien - Stage 2 models

First Lien-Stage 2: Partial loss rate	
Variable	**Variance Inflation**
b_Intercept	0
b_log_curr_bal	1.36528
b_LTV_MTM_SP_gt90	3.46141
b_pmi	1.01878
b_judicial	1.09392
b_BROK_IND	1.06881
b_prin_reduction	1.00105
b_UnempRate	4.29802
b_log_HPI_t_orig_ratio	1.44561

Second Lien-Stage 2: Partial Loss rate	
Variable	**Variance Inflation**
b_Intercept	0
b_log_fst_mtg_bal	1.12771
b_cLTV_MTM_SP_gt80_le120	1.22415
b_UnempRate	1.50377
b_deficiency	1.34921
b_CBNA_ind	1.00914
b_REL_ind	1.06576

6.1.4 Stationarity

To the extent applicable, provide evidence that the time series data are stationary. [Only applicable for time series models]

> [List the tests of stationarity conducted. For the test requirements refer to the Model Testing Guidance. Minimum requirement for econometric time series model is to test stationarity of the model residuals. To the extent applicable, test also stationarity of both the dependent and selected independent time series of the model. Comment on the results of the tests. If not tested, please explain.]

Per MRM model performance testing guidance (issued June 2018) this is not applicable, as the 2019 Method A PD and LGD models are not time series models. Per MRM model performance testing guidance (issued June 2018), this is not applicable, as the 2019 Method A PD and LGD models are not time series models. Please note macro-econ variables such as HPI growth rate and unemployment rate are not modeled (i.e. dependent) variables. Instead, they have been provided by GCRM as standard macro-econ inputs for all CCAR models. Second, the transition model itself is not time series model. Therefore, stationary test is not applicable here. Last but not the least, the logistic regression used by the transition model does not require its input variables to be stationary.

Question - Please provide the stationarity test results for the lagged/transformed macro-variables used in the final model equations.

Answer – As requested by MRM, CAMU conducted stationarity tests on relevant lagged/transformed dependent attributes. Please refer to attachment - - 6.1.4 Stationarity Results.doc.' for details.

6.1.5 Autocorrelation

> To the extent applicable, provide evidence that residuals are not auto-correlated. [Only applicable for time series models.]
>
> [List the tests of autocorrelation conducted. For the test requirements refer to the Model Testing Guidance. Comment on the results of the tests. If not tested, please explain.]

Per MRM model performance testing guidance (issued June 2018) this is not applicable, as the 2019 Method A PD and LGD models are not time series models.

6.1.6 Normality

To the extent possible, provide the evidence that the residuals are normally distributed

[List the tests of normality conducted. For the test requirements refer to the Model Testing Guidance. Comment on the results of the tests. If not tested, please explain.]

Per MRM model performance testing guidance, this is not applicable, as the 2019 Method A PD models use logistic regressions and the LGD model uses nonlinear regression.

6.1.7 Heteroscedasticity

To the extent applicable, provide evidence that there is no heteroscedasticity of residuals

[List the tests of heteroscedasticity conducted. For the test requirements refer to the Model Testing Guidance. Comment on the results of the tests. If not tested, please explain.]

Per MRM model performance testing guidance, this is not applicable, as the 2019 Method A PD models use logistic regressions and LGD model uses nonlinear regression.

6.1.8 Model Goodness of Fit

Predictive accuracy of the model

Provide evidence that the predictive accuracy of the model is sufficiently high. Evaluate the accuracy and consistency of the model.

[Evaluate the performance measures as per the Model Testing Guidance given the model methodology and individual model components. Comment on the resulting value of performance measures.]

Predictive accuracy of the models has been tested by analyzing the trends of actual and predicted values over the development sample. For each equation in Method B PD models, CAMU evaluates the predictive accuracy of the model mainly based on the predicted error, RMSE, COV and AIC statistics, the definitions of which are discussed below.

The predicted error is calculated by:

$$PredError = \frac{1}{T}\sum_{t=1}^{T}\frac{\hat{y}_t - y_t}{y_t}$$

where T is the number of observation months in development data, and y_t and $\hat{y}_t$ are the mean of actual roll rate and mean of predicted roll rate, respectively, at observation month t.

The RMSE is calculated by:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[(\hat{y}_t - y_t)^2]}$$

COV is calculated by

$$COV = RMSE / \frac{1}{T}\sum_{t=1}^{T} y_t$$

RMSE and COV measures how much the predicted values deviate from the actuals. Low error and low COV are statistically desirable for a better fit.

The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given set of data. AIC quantifies 1) the goodness of fit (predictive accuracy), and 2) the simplicity/parsimony, of the model into a single statistic.

Table 6.1.8.1 presents the predicted error, root mean square error (RMSE), coefficient of variation (COV), AIC with intercept only, and AIC with intercept and covariates for all 28 Residential Mortgage PD equations estimated in the system of complete transition matrix. All the statistic metrics were calculated on model development sample only. Please refer to tab 'Goodness of Fit' of the attachment – '6.1 Residential Mortgage Model Diagnostic and Statistical Tests.xlsx'.

The error ranges from -0.190 to 0.185, RMSE ranges from 0.002 to 0.040, COV ranges from 0.072 to 0.942 for the PD model. Most of transition equations show reasonably small error, RMSE and COV, demonstrating high predictive accuracy of the model. The high predictive accuracy is further confirmed by the robust overall model fitting and robust back test results as discussed in section 6.3 Back-Testing. As shown below, in a couple of equations modelers observed relatively elevated errors and/or RMSE (e.g., BUK7->VP, BUK4->IVP, BUK5->IVP, BUK6->BUK4, BUK5->BUK01, BUK6->BUK01 and BUK7->BUK01), but all of instances have very small population sizes or low transition rates, so they have a minimal impact to the overall model performance. Nevertheless, as shown by AICs, adding covariates improved the model performance compared to a simple average. Therefore, modeling these transitions can still improve overall model predictive accuracy and increase model sensitivity. Further discussion can be referred in section 6.1.9 one-month-ahead analysis.

Table 6.1.8.2 - 6.1.8.5 presents the predicted error, root mean square error(RMSE), COV, AIC with intercept only, and AIC with intercept and covariates, R-Squared, C-statistics and other model predictive accuracy tests results for the LGD model. All the statistic metrics were calculated on model development sample only and are satisfactory, hence validating the LGD model predictive accuracy. Please refer to tab '6.1.8 Goodness of Fit' of the attachment – '6.1 Severity Model Diagnostic and Statistical Tests'.xlsx.

Table 6.1.8.1 Predicted error, RMSE and AIC for all residential mortgage PD equations

Model	Pred Error	Pred RMSE	COV	AIC: Intercept Only	AIC: Intercept and Covariates
BUK01->BUK2	0.004	0.002	0.099	8,067,897	6,545,026
BUK01->VP	0.028	0.002	0.155	4,747,931	4,520,027
BUK2->BUK01	0.029	0.027	0.072	1,863,379	1,803,939
BUK2->BUK3	0.025	0.026	0.085	1,589,875	1,540,273
BUK2->VP	0.063	0.007	0.285	192,721	163,699
BUK3->BUK01	0.057	0.038	0.165	372,388	348,338
BUK3->BUK2	0.020	0.032	0.103	455,186	428,603
BUK3->BUK3	0.006	0.040	0.078	694,664	666,451
BUK3->VP	0.017	0.008	0.319	53,250	44,631
BUK4->BUK01	0.075	0.021	0.198	154,857	146,596
BUK4->BUK2	0.028	0.018	0.248	106,449	100,683
BUK4->BUK3	0.003	0.025	0.138	217,065	206,135
BUK4->BUK4	0.025	0.033	0.088	402,930	379,582
BUK4->IVP	-0.190	0.004	0.647	17,394	17,072
BUK5->BUK01	0.129	0.020	0.358	93,904	89,095
BUK5->BUK2	0.071	0.007	0.391	35,014	33,924
BUK5->BUK3	0.069	0.009	0.313	53,049	51,243
BUK5->BUK4	0.082	0.020	0.234	127,066	122,485
BUK5->BUK5	0.048	0.033	0.157	270,315	258,920
BUK5->IVP	-0.077	0.011	0.841	34,425	30,828
BUK6->BUK01	0.110	0.022	0.422	70,905	67,923
BUK6->BUK4	0.185	0.010	0.529	35,432	34,649
BUK6->BUK5	0.001	0.024	0.429	85,480	82,750
BUK6->BUK6	0.086	0.038	0.231	197,517	191,124
BUK6->IVP	-0.086	0.020	0.245	140,759	75,321
BUK7->BUK01	0.103	0.006	0.326	302,121	286,507
BUK7->IVP	0.025	0.010	0.203	715,951	686,964
BUK7->VP	0.150	0.005	0.942	108,738	102,890

Table 6.1.8.2: LGD Performance measures results

Model component/ segment	Type of test	Test results	Business justification/ Limitations	Mitigation
First Lien / Stage1	C-Statistics, AIC, MAD, RMSE, COV	Satisfactory	No major limitation	N/A
First Lien/ Stage 2	R-squared, RMSE	Satisfactory	No major limitation	N/A
Second Lien / Stage1	C-Statistics, AIC, MAD, RMSE, COV	Satisfactory	No major limitation	N/A
Second Lien/ Stage 2	R-squared, MAD, RMSE, COV	Satisfactory	No major limitation	N/A

Table 6.1.8.3: Akaike information criterion (AIC) of stage 1 models

Portfolio	Equation	AIC	Likelihood

		Intercept Only	Intercept and Covariates	Ratio
First Lien	Zero Vs Partial	14,691	13,653	1,042
	Full Vs Partial	48,719	36,790	11,943
Second Lien	Zero Vs Partial	11,537	11,073	470
	Full Vs Partial	105,074	95,854	9,239

Table 6.1.8.4: R-Squared of Stage 2 models

Portfolio	R-Square
First Lien	86%
Second Lien	94%

Table 6.1.8.5: Model Predictive Accuracy Test

Portfolio	Equation	MAD	RMSE	COV
First Lien	Stage 1: Zero Vs Partial	0.01	0.02	0.54
	Stage 1: Full Vs Partial	0.02	0.03	0.28
	Stage 2: Partial Loss Rate	0.02	0.03	0.07
Second Lien	Stage 1: Zero Vs Partial	0.03	0.04	0.61
	Stage 1: Full Vs Partial	0.03	0.03	0.04
	Stage 2: Partial Loss Rate	0.03	0.05	0.06

Discriminatory power of the model

To the extent applicable, provide evidence that the discriminatory power is sufficiently high. Evaluate the model's ability to discriminate between good/bad or rank order outcomes.

[Evaluate the performance measures as per the Model Testing Guidance given the model methodology and individual model components. Comment on the resulting value of performance measures].

The discriminatory power of the RM model is evaluated using C statistics. It is derived from Receiver Operating Characteristics (ROC) curve, in which the true positive rate is plotted against the false positive rate and indicates discrimination power of model. The C-statistics is directly implemented in the logistic regressions of SAS "Proc Logistic", enabling computational efficiency. Therefore, C-statistics is used to assess the discriminatory power of Method A Residential Mortgage PD/LGD models.

Table 6.1.8.6 below shows C-statistics for all 28 PD transition equations. All regressions have achieved relatively high value in C-statistics and meet MRM required threshold of 0.6, indicating a consistently strong discrimination power. The summary table of model goodness of fit can also be accessed in Goodness of fit tab in the attached excel 6.1 Residential Mortgage Model Diagnostic and Statistical Tests.xlsx.

Similar to PD, C statistics have been used to assess the discriminatory power of first lien and second lien stage 1 LGD models. As before, a relatively high value of C-statistics indicates a strong discrimination power for stage 1 models as shown in Table 6.1.8.7. Please refer to tab '6.1.8 Goodness of Fit' of the attachment – '6.1 Severity Model Diagnostic and Statistical Tests'.xlsx.

Table 6.1.8.6: C-Statistics for all residential mortgage PD equations

Model	C-statistics
BUK01->BUK2	0.853
BUK01->VP	0.681
BUK2->BUK01	0.619
BUK2->BUK3	0.618
BUK2->VP	0.809
BUK3->BUK01	0.676
BUK3->BUK2	0.66
BUK3->BUK3	0.629
BUK3->VP	0.833
BUK4->BUK01	0.676
BUK4->BUK2	0.687
BUK4->BUK3	0.662
BUK4->BUK4	0.661
BUK4->IVP	0.649
BUK5->BUK01	0.684
BUK5->BUK2	0.663
BUK5->BUK3	0.659
BUK5->BUK4	0.644
BUK5->BUK5	0.635
BUK5->IVP	0.795
BUK6->BUK01	0.668
BUK6->BUK4	0.626
BUK6->BUK5	0.641
BUK6->BUK6	0.621
BUK6->IVP	0.916
BUK7->BUK01	0.699
BUK7->IVP	0.656
BUK7->VP	0.714

Table 6.1.8.7: C-Statistics of Stage 1 models

Portfolio	Equation	C-Statistics
First Lien	Stage 1: Zero Vs Partial	80%
	Stage 1: Full Vs Partial	83%
Second Lien	Stage 1: Zero Vs Partial	68%
	Stage 1: Full Vs Partial	65%

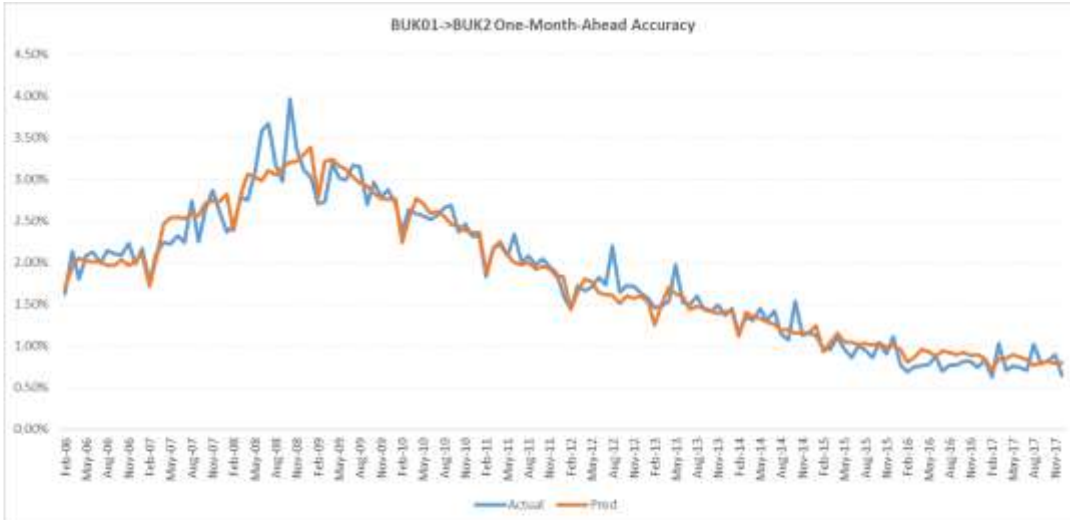
6.1.9 Other tests – Model-specific diagnostics

- Provide evidence of other relevant tests performed. Examples include linearity in the context of typical econometric regression models.

An additional model diagnostic is one-month ahead analysis for PD model. Although the Residential Mortgage PD model was developed based on a binary logistic structure, CAMU also evaluated each individual model's one-month ahead performance by converting binary logistic results into a multinomial framework. This assessment was undertaken to evaluate how the model was performed for the conditional transitions over time.

For the most important transitions representing relatively large population sizes of the portfolio or representing large sources for default, CAMU has observed reasonably satisfactory model performance. Figure 6.1.9.1 displayed the one-month-ahead model performance for the most important delinquent transition BUK01->BUK2 as an example. Details for each PD model can be found in the attachment excel file ‘6.1.9 Residential Mortgage PD One Month Ahead.xlsx.’

Figure 6.1.9.1 One-month-ahead model performance for BUK01->BUK2



- To the extent that there are Model Assumptions listed in Section 5.3 that are not evaluated/tested elsewhere in Section 5, list in this sub-section

  [Explain what other model specific diagnostics were performed. For the test requirements refer to the Model Testing Guidance. Comment on the results of these other tests.]

  Please refer to Section 5.3 for all Model Assumptions. Specifically for the non-modeled assumptions, please refer to Section 6.4.4.

- Specific to Vendor Models, apply an appropriate set of validation tests to determine model performance and robustness for Bank's portfolio.

  This is Not Applicable. The 2019 Method A model suite has been internally developed within CAMU, so they are not vendor models.

- For vendor models, describe if additional tests are performed or additional benchmarks used to gauge the model performance, and indicate if the model output is reasonable.

  This is not applicable. The 2019 Method A model suite has been internally developed within CAMU, so they are not vendor models.

6.2 Model Robustness and Stability

The aim of this section is to assess the performance of the model. The results of model stability testing must be presented in this section. Please refer to Model Testing Guidance for detailed requirements.

In the modeling process, predicted error and RMSE were used to evaluate the model's predictive accuracy. AIC was also used as a measure of performance. C-statistics were used to evaluate model discriminatory power. These metrics have been presented and discussed in section 6.1.8. In addition, CAMU evaluated each individual model's one-month-ahead performance in a multinomial framework which has been discussed in section 6.1.9. Furthermore, CAMU conducted model validation on an in-time out-of-sample and full sample which includes both in-time and out-of-time data for each individual model robustness and stability tests. Details will be elaborately discussed in section 6.2 later.

Below are the key observations and evidence to support 2019 Method A Residential Mortgage Model strong performance:

- Predicted errors for key transitions are within a satisfactory range.
- Consistently high C-statistics were observed, thereby indicating that most models displayed strong discriminatory power.
- AICs for all models were smaller compared to intercept only models indicating that models were more accurate relative to intercept only models.
- Strong and robust model performance are further supported by the comparison between monthly predicted and actual transition probabilities in one month ahead model output.
- Parameters were consistently significant, showed business intuitive sign and did not vary largely across different samples, which indicates a fair model stability.

6.2.1 Model Stability Design

[Please explain the Short Sample/ Out of Sample (OOS)/ Out of Time (OOT) design. Clearly specify the OOS/OOT and the development time period. In case of short sample clearly specify the numbers of months excluded from the development sample to derive the OOS period. Refer to Model Testing Guidance for requirement on Model Stability principles and testing].

The 2019 Method A modeling team conducted model validation on the in-time (INT) hold-out sample and full sample at equation level to comply with MRM testing guidance to ensure parameter stability across development and validation samples.

Specifically for the 2019 Method A Residential Mortgage PD model, data from Feb 2006 to Dec 2011 and from Apr 2014 to Dec 2017 is considered as in-time sample, of which 80 percent is used as model development data and the remaining 20 percent is used as INT validation data. The data from Jan 2012 to Mar 2014 (27 months) is considered as out-of-time data.

For the 2019 Method A LGD model, data from Jan 2008-to Dec 2011 and from Apr 2014-to Jun 2017 is considered as in-time sample, of which 80% percent is used as model development data and the remaining 20% percent is used as INT validation data. The data from Jan 2012- Mar 2014 (27 months) is considered as out-of-time data.

The full sample contains development data, in-time hold out data as well as out-of-time data. The rationale for designing the samples in this way is to ensure that the development data can include both stress period and most recent period data.

Samples collected between Jan 2012 and Mar 2014 was chosen to be the out-of-time validation sample, whereas the model development sample include both observations collected on the stress period and most recent period. CAMU believes the choice of the observation window has the following three advantages:

1) It enables the model to be developed on a wider spectrum of economic conditions and loan origination credit policy.

2) It includes more loans originated after the crisis so that the model development data can better resembles the go-forward portfolio as of today.

3) It utilizes more loans that have gone through the ARM/IO reset process so the model can better capture the reset risk associated with the portfolio

For each individual model equation, the parameters are estimated on model development data and then compared across INT and full samples. Variables survived in the final model equation has to meet the following three criteria.

iv. Parameter sign remains the same across development data, INT, and full sample
v. Parameter estimates remain statistically significant (P-value <= 0.05) across development data, INT, and full sample
vi. Magnitude of parameter estimate do not have largely shift across development data, INT, and full sample

Detailed results have been described in MDD section 6.2.2.

The stability analysis is conducted among DEV/in-time/Long sample. The Long sample includes both the model development period and the OOT periods. Yes, the estimates from the 80% DEV data are used as the final model specifications. Please note, Long sample is same as Full sample. They have the same connotation and have been interchangeably used.

6.2.2 **Model Stability Results**

How was model robustness and stability ensured and tested?

[Evaluate the stability of the model in terms of Shifts in magnitude, Coefficient sign and Significance of coefficients as per the Model Testing Guidance. Provide the results for Short Sample/ Out of Sample (OOS)/Out of Time (OOT) results as applicable.]

For 2019 Method A Residential Mortgage PD/LGD models, modelers performed INT and full sample validation testing on each individual model equation, providing solid evidence that all 28 PD transition models and the 6 LGD stage 1 and stage 2 models have satisfactorily demonstrated stability and robust across different samples.

Table 6.2.2.1 below shows the PD model estimation results across different samples for model BUK01->BUK2. For the complete list of PD model stability results, please refer to the attached **6.2 Residential Mortgage Model Robustness and Stability Results.xlsx.**

Tables 6.2.2.2a and 6.2.2.2b shows the LGD model estimation results comparison between development and Long samples and the OOS and Development samples respectively for the first lien loans. **Tables 6.2.2.3a and 6.2.2.3b** shows the LGD model estimation results comparison between full vs in-time sample vs OOT samples respectively for the second lien loans. As shown in the listed tables, all variables remain statistically significant and keep business intuitive sign across three samples. Parameter shift measures the parameter magnitude change compared with model development sample. No large shifts were detected for key risk drivers. **Tables 6.2.2.4 and 6.2.2.5** lists the C-statistics comparisons between Full vs INT vs OOT samples respectively for the Stage 1 model and R-squared comparisons for the Stage 2 model. For the complete list of LGD model stability results, please refer to the attached **6.2 Severity Model Robustness and Stability.xlsx.**

Table 6.2.2.1 BUK01->BUK2 model estimation for DEV vs. OOS vs. Full samples

Parameter	DEV		INT			Full		
	Estimate	Pr > ChiSq	Estimate	Pr > ChiSq	Parameter shift(%) INT vs DEV	Estimate	Pr > ChiSq	Parameter shift(%) Full vs DEV
Intercept	2.954	<.0001	3.572	<.0001	21%	3.261	<.0001	10%
B_S_BROK_IND	0.322	<.0001	0.340	<.0001	5%	0.369	<.0001	15%
B_S_CORR_IND	0.554	<.0001	0.543	<.0001	-2%	0.556	<.0001	0%
B_S_WLST_IND	0.690	<.0001	0.636	<.0001	-8%	0.679	<.0001	-2%
B_S_LoanPurpPurch_IND	0.049	<.0001	0.061	<.0001	24%	0.061	<.0001	25%
B_S_LoanPurpWork_IND	0.922	<.0001	1.115	<.0001	21%	0.954	<.0001	3%
B_S_OCC_IV_IND	0.315	<.0001	0.297	<.0001	-6%	0.290	<.0001	-8%
B_S_lowIncDoc_IND	0.303	<.0001	0.348	<.0001	15%	0.286	<.0001	-5%
B_S_pmi_IND	0.159	<.0001	0.095	<.0001	-40%	0.156	<.0001	-2%
B_S_DocMISS_IND	0.158	<.0001	0.136	<.0001	-14%	0.145	<.0001	-8%
B_S_LoanPurpMiss_IND	0.401	<.0001	0.408	<.0001	2%	0.407	<.0001	1%
B_M_coborrower_IND	-0.245	<.0001	-0.220	<.0001	-10%	-0.245	<.0001	0%
ARM_5Yr	-0.267	<.0001	-0.304	<.0001	14%	-0.262	<.0001	-2%
ARM_GT5Yr	-0.318	<.0001	-0.415	<.0001	30%	-0.311	<.0001	-2%
P_S_OrigIntSpread_FRM	0.081	<.0001	0.074	<.0001	-9%	0.079	<.0001	-3%
B_S_INT_ONLY_IND	-1.215	<.0001	-1.248	<.0001	3%	-1.227	<.0001	1%
POST_IO_IND	0.148	<.0001	0.222	<.0001	50%	0.159	<.0001	8%
P_AND_I_ratio_IO	0.600	<.0001	0.720	<.0001	20%	0.634	<.0001	6%
POST_ARM_6M_IND	0.198	<.0001	0.208	<.0001	5%	0.202	<.0001	2%
POST_ARM_IND	-0.081	<.0001	-0.105	<.0001	31%	-0.094	<.0001	16%
D_M_PRIN_BAL_LE2K	0.969	<.0001	0.920	<.0001	-5%	0.983	<.0001	1%
Balloon_Mature	1.458	<.0001	1.764	<.0001	21%	1.379	<.0001	-5%
Katrina	0.098	<.0001	0.102	0.0003	4%	0.099	<.0001	1%
Sandy	0.212	<.0001	0.376	<.0001	77%	0.279	<.0001	31%
FEB	-0.248	<.0001	-0.235	<.0001	-6%	-0.241	<.0001	-3%
MAR	-0.098	<.0001	-0.112	<.0001	15%	-0.104	<.0001	6%
S_M_FicoRefresh	-0.010	<.0001	-0.010	<.0001	0%	-0.010	<.0001	-1%
S_M_FicoRefresh_SP640	-0.010	<.0001	-0.010	<.0001	-1%	-0.010	<.0001	-2%
S_M_FicoRefresh_SP720	0.009	<.0001	0.009	<.0001	-6%	0.009	<.0001	-5%
B_S_missS_M_FicoRefresh	0.240	<.0001	0.255	<.0001	6%	0.264	<.0001	10%
N_M_age_oldest_trd	-0.001	<.0001	-0.001	<.0001	0%	-0.001	<.0001	-4%
miss_N_M_age_oldest_trd	0.065	<.0001	0.100	<.0001	54%	0.064	<.0001	0%
Post2010_Orig_CONV	-0.566	<.0001	-0.501	<.0001	-12%	-0.570	<.0001	1%
HPI_tm12_ratio_LSP1_CMI_CONV	-0.552	<.0001	-0.474	<.0001	-14%	-0.532	<.0001	-4%
MTM_CLTV_SP40_1st_CONV	0.004	<.0001	0.003	<.0001	-21%	0.004	<.0001	-7%
MTM_CLTV_SP80_1st_CONV	0.003	<.0001	0.005	<.0001	37%	0.004	<.0001	13%
MTM_CLTV_SP80_2nd_CONV	0.005	<.0001	0.006	<.0001	9%	0.006	<.0001	3%
post_neg_CONV	-0.214	<.0001	-0.229	<.0001	7%	-0.227	<.0001	6%
R_M_State_UnempB12M_IO_CONV	0.655	<.0001	0.602	<.0001	-8%	0.630	<.0001	-4%
R_M_State_UnempB12M_BROK_CONV	0.107	<.0001	0.058	0.0036	-46%	0.065	<.0001	-39%
income_12m_ratio	-0.576	<.0001	-0.935	<.0001	62%	-0.707	<.0001	23%
N_M_MOB_cap_WLST_CONV	-0.002	<.0001	-0.002	<.0001	-13%	-0.002	<.0001	-8%
P_M_CURR_NOTE_RATE_ARM_CONV	0.071	<.0001	0.070	<.0001	-1%	0.071	<.0001	0%
D_M_PRIN_BAL_in10K_CMI_CONV	0.002	<.0001	0.002	<.0001	-9%	0.002	<.0001	-8%
P_M_State_Unemp_Rate_lag12_GOV	0.026	<.0001	0.022	<.0001	-14%	0.033	<.0001	30%
MTM_CLTV_SP80_GOVGT3	0.009	<.0001	0.011	<.0001	19%	0.009	<.0001	1%

Table 6.2.2.2a: Model Stability test of LGD - First Lien for Full vs INT Models

First Lien-Stage 1: Zero Vs Partial			
Variable	**Statistical Significance change**	**Coefficient Sign change**	**Coefficient Magnitude change%**
Intercept	No	No	-8%

LTV_mtm_SP_gt20_le100	No	No	1%
log_HPI_tm12_ratio	No	No	-21%

First Lien-Stage 1: Full Vs Partial			
Variable	**Statistical Significance change**	**Coefficient Sign change**	**Coefficient Magnitude change%**
Intercept	No	No	4%
log_curr_bal_SP_le300k	No	No	-3%
log_curr_bal_SP_gt300k	No	No	6%
log_HPI_tm12_ratio_SP_le0	No	No	-40%
UnempRate_SP_gt9	No	No	-10%
Judicial	No	No	1%
HLC_ind	No	No	-7%
FHL_ind	No	No	-3%

First Lien-Stage 2: Partial loss rate			
Variable	**Statistical Significance change**	**Coefficient Sign change**	**Coefficient Magnitude change%**
b_Intercept	No	No	15%
b_log_curr_bal	No	No	8%
b_LTV_MTM_SP_gt90	No	No	-2%
b_pmi	No	No	-6%
b_judicial	No	No	7%
b_BROK_IND	No	No	0%
b_prin_reduction	No	No	4%
b_UnempRate	No	No	-8%
b_log_HPI_t_orig_ratio	No	No	14%

Table 6.2.2.2b: Model Stability test of LGD First Lien for OOS vs INT Models

First Lien-Stage 1: Zero Vs Partial			
Variable	**Statistical Significance change**	**Coefficient Sign change**	**Coefficient Magnitude change%**
Intercept	No	No	43%
LTV_mtm_SP_gt20_le100	No	No	-4%
log_HPI_tm12_ratio	No	No	-38%

First Lien-Stage 1: Full Vs Partial

Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
Intercept	No	No	3%
log_curr_bal_SP_le300k	No	No	7%
log_curr_bal_SP_gt300k	No	No	1%
log_HPI_tm12_ratio_SP_le0	No	No	7%
UnempRate_SP_gt9	No	No	2%
Judicial	No	No	-14%
HLC_ind	No	No	4%
FHL_ind	No	No	3%

First Lien-Stage 2: Partial loss rate			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
b_Intercept	No	No	19%
b_log_curr_bal	No	No	9%
b_LTV_MTM_SP_gt90	No	No	-2%
b_pmi	No	No	-3%
b_judicial	No	No	8%
b_BROK_IND	No	No	24%
b_prin_reduction	No	No	26%
b_UnempRate	No	No	-31%
b_log_HPI_t_orig_ratio	No	No	52%

Table 6.2.2.3a: Model stability test of LGD Second Lien for Full vs INT Models

Second Lien-Stage 1: Zero Vs Partial			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
Intercept	No	No	-2%
cLTV_MTM_unemp_down	No	No	-8%
REL_ind	No	No	-5%
HLC_ind	No	No	-13%

Second Lien-Stage 1: Full Vs Partial			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
Intercept	No	No	-3%
log_curr_bal_SP_gt10k_le150k	No	No	11%
log_curr_bal_SP_gt150k	No	No	-2%
junior_ratio_SP_lep20	No	No	1%

log_HPI_tm12_ratio	No	No	12%
R_M_UnempB12M_SP_gt1_le1p2	No	No	9%
Lpi	No	No	3%
Judicial	No	No	12%
wasUSRB	No	No	2%
REL_ind	No	No	21%
prin_reduction	No	No	-1%

Second Lien-Stage 2: Partial Loss rate			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
b_Intercept	No	No	4%
b_log_fst_mtg_bal	No	No	-2%
b_cLTV_MTM_SP_gt80_le120	No	No	35%
b_UnempRate	No	No	-10%
b_deficiency	No	No	11%
b_CBNA_ind	No	No	10%
b_REL_ind	No	No	10%

Table 6.2.2.3b: Model stability test of LGD Second Lien for OOS vs INT Models

Second Lien-Stage 1: Zero Vs Partial			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
Intercept	No	No	-1%
cLTV_MTM_unemp_down	No	No	3%
REL_ind	No	No	-7%
HLC_ind	No	No	-5%

Second Lien-Stage 1: Full Vs Partial			
Variable	Statistical Significance change	Coefficient Sign change	Coefficient Magnitude change%
Intercept	No	No	12%
log_curr_bal_SP_gt10k_le150k	No	No	-3%
log_curr_bal_SP_gt150k	No	No	14%
junior_ratio_SP_lep20	No	No	9%
log_HPI_tm12_ratio	No	No	3%
R_M_UnempB12M_SP_gt1_le1p2	No	No	16%
lpi	No	No	6%
judicial	No	No	-10%
wasUSRB	No	No	-15%
REL_ind	No	No	76%
prin_reduction	No	No	-2%

Second Lien-Stage 2: Partial Loss rate			
Variable	**Statistical Significance change**	**Coefficient Sign change**	**Coefficient Magnitude change%**
b_Intercept	No	No	0%
b_log_fst_mtg_bal	No	No	0%
b_cLTV_MTM_SP_gt80_le120	No	No	-6%
b_UnempRate	No	No	6%
b_deficiency	No	No	12%
b_CBNA_ind	No	No	-9%
b_REL_ind	No	No	2%

Table 6.2.2.4: LGD C-Statistics for Full vs INT vs OOS Samples – Stage 1 Model

Portfolio	Equation	C-Statistics of Stage 1/Dev	C-Statistics of Stage 1/OOS	C-Statistics of Stage 1/All
First Lien	Stage 1: Zero Vs Partial	80%	80%	82%
	Stage 1: Full Vs Partial	83%	84%	81%
Second Lien	Stage 1: Zero Vs Partial	68%	67%	67%
	Stage 1: Full Vs Partial	65%	66%	67%

Table 6.2.2.5: LGD R-Squared for Full vs INT vs OOS Samples – Stage 2 Model

Portfolio	R-Square of Stage 2/DEV	R-Square of Stage 2/OOS	R-Square of Stage 2/All
First Lien	86%	86%	86%
Second Lien	94%	94%	94%

- Are the Model Stability results satisfactory? If not provide justification

[If the Stability results were not observed to be satisfactory in terms of shifts in the magnitude, coefficients sign and significance of coefficient please provide the rationale for use of the variable in the model].

As described above, INT and full sample validations for 2019 Method A Residential Mortgage models showed that all parameter estimates remain statistically significant, all parameter signs remain the same, and all parameter magnitudes are relatively stable.

6.3 Back-Testing

[Back-testing on historical model performance should be carried out, highlighting any known conditions where the model underperformed. For PPNR and Retail Loss forecasting models, tests on stress, non-stress and recent observed periods must be performed as applicable to the model usage. Clearly mention the backtest period for each of the tests. The back-testing window should be in line with the model usage (CCAR, DFAST, ALLL, IFRS9, CECL etc.) and model forecast horizon (short/medium/long term). The backtest results should be in line with the model usage grid].

Back-testing principles and a minimum set of back-tests that should be performed can be found in the Model Testing Guidance attached at the beginning of section 8. All the back-testing results should be provided in the Back-testing Template which is embedded in the Model Testing Guidance.]

A predictive model's back testing fulfils five major goals –

1. The first goal is to ensure that the model performs with sufficient accuracy in both stress and non-stress environments. Back-testing projected results versus actual results provide actionable feedback on the validity of the forecast.

2. Backtest results also provide additional indirect confirmation regarding model stability and robustness over time.

3. It provides a rigorous method to test key components of the model and to identify the size of potential errors over different periods, as stipulated by the Model testing Guidance Requirements.

4. It satisfies the regulatory mandates (CCAR and Dodd-Frank) which lists model backtests as an essential part of the model validation process

5. Provides a holistic assessment of the model performance attributed to the portfolio composition changes over time

To achieve the afore-mentioned goals, back testing is conducted on the Method A RM model using portfolio snapshots over specified time horizons measuring the model's performance through the historical stress period and more recent periods.

The Method A RM Model backtesting indicates a robust predictive model with reasonably strong performance. For a range of portfolio snapshots and market environments, the model accurately predicts the most important output metrics. Error rates are all well below the +/-25% threshold as established by MRM. Furthermore, error rates among the delinquency buckets and IVP / VP terminal states (i.e., loan transition statuses) are all low, indicating strong prediction capability for the evolution of loan status.

Further, the characteristic analysis provides additional leverage to the model's chosen segmentation scheme. It demonstrates that the model appropriately captures key risk drivers that affect the model performance. Typically, credit profiles are sensitive to FICO, LTV, and delinquency statuses. The model appropriately risk-ranks FICO bands, LTV bands, and statuses, with error rates of reasonable magnitude.

6.3.1 Back-testing design

[Describe the design of the back-tests performed. If any recommended tests are omitted, explain the reasons and list model performance limitations due to missing back-tests. For all recommended back-tests performed, describe the data sample and model components that were used. If additional back-tests are performed, describe and explain the model component(s) tested, the data used, the assumptions that were applied, the business rationale for these tests, and the formulas used to evaluate the back-test results.]

The Method A RM Model is a suite of models: the PD model which is reflective of a loan-level transition structure yielding a monthly transition matrix, EAD which is based on amortization schedule and the LGD model which models the losses based on their outcome(full, zero, partial). In particular, the individual components(or equations) of the PD model in combination with the EAD amortization schedule, the DV Logic and the LGD model work together in a unified system to yield economically important metrics which produces the model end-to-end results. These behavioral model results, combined with business logic/algorithms, result in forecasted voluntary prepay (henceforth referred to as VP), involuntary prepay (henceforth referred to as IVP), intermediate delinquency bucket projections and losses (gross and net). These metrics discussed above are deemed as the relevant metrics, as they summarize the model's complex interactions, leading to the outcomes, which impact business performance, and are the main outputs deemed appropriate for the mandated CCAR reporting.

Presented within this section are the model component level and end-to-end results at the overall portfolio level and the relevant segment(s) level. Although equation-level model performance is important and is provided within the supporting files and attachments, integrated model performance is more critical.

Backtesting performance is thus assessed via the following metrics:

- GCL (AMT): the key outcome variable from the PD model, incorporating the influence of the transition matrix and EAD amortization. GCL is calculated as first time 180+ days past due (DPD) and direct IVP (IVP from lower than 180DPD) for never 180DPD loans, and IVP forever 180DPD loans
- GCL (Units): represents the summation of the loan count which are first time 180+ days past due (DPD) and direct IVP (IVP from lower than 180DPD) for never 180DPD loans, and IVP forever 180DPD loans
- GCL%ENR: It is calculated as the GCL amount at time t divided by ENR at time t.
- NCL (AMT): Net Credit Loss is the key outcome variable from both PD and LGD model, incorporating the influence of all models and modelled variables.
- NCL%ENR: It is calculated as the NCL amount at time t divided by ENR at time t.
- Delinquency Buckets: Reflects the dollar amount and units exposure at each delinquency bucket
- VP: Reflects the terminal status balances and units for loans that voluntarily prepaid

- EAD: EAD(t) = EAD(t-1) – Forecasted Payment Amount of Principal Balance(t) and EAD(0) = UPB(0)
- VP: Reflects the terminal status balances and units for loans that voluntarily prepaid

For the Annual CCAR Capital Adequacy Assessment (CCAR) and Mid Cycle Stress testing (MCST) requirements, the recommended model outputs correspond to the delinquency bucket and NCL projections. The other outputs caters to the ALLL and Repurchase reserve estimates, CECL Reporting and other BAU model usages as stated in Section 1.1 within the Usage Grid.

The backtesting methodology leverages the naive censoring approach for effective modeling of the action loans, as recommended by MRM. One important point to note here is that 'Censoring' has no effect on model forecasting results or the model's sensitivity to macro-economic factors as the model forecast always assumes that there are no future 'action' loans. Under naïve censoring, the model removes both actual and predicted post-action performance for action loans.

The back-test assesses the model's ability to accurately project future defaults and losses arising from the snapshot population throughout the chosen forecast horizon. Actual outcomes are compared to the forecasted output from the model, which helps to establish the robustness of the end-to-end model suite.

- Has back-testing been conducted at various levels of the model - at the model component level (for e.g. PD, LGD, EAD), as well as final output (for e.g., NCL, Balance, Revenue) and as per the model usage grid?

  [Provide a scheme outlining the model structure and its components. Highlight what was tested and explain why certain components might have been omitted, e.g. because of assumptions.]

Yes, CAMU conducted backtesting on multiple levels commensurate with MRM's model performance testing guidance and the Model Usage Grid. The GCL-centric variables (GCL AMT, Unit and GCL%ENR) reflect the PD model performance which incorporates the EAD amortization schedule. The LGD model leverages the PD model outputs as its inputs and generates estimates of NCL(AMT and NCL%ENR) which effectively is a reflection of the model's end-to-end results. Further the error differential between the GCL-centric and NCL-centric metrics implicitly backtests the LGD model as GCL times LGD defines net credit loss (NCL).

$$NCL = GCL * LGD$$
$$LGD = \frac{NCL}{GCL}$$

It is important to note here that NCL and GCL are not contemporaneous events. NCL is the net credit loss happened during a period of time. It is not "finalized" until recoveries are completed, which could take time. GCL, however, is mainly determined by 180DPD event. For loans already hit 180DPD, IVP is used to measure GCL. The subtle difference in GCL and NCL, especially the timing mismatch potentially affects the results in the backtesting horizon.

- Has back-testing been conducted on different time periods – recent, stress period, non-stress period (if applicable)? Do the types of back-tests conducted sufficiently cover all potential model risks within its performance and use?
  [List all the back-tests conducted. Describe the data used, and link it to the components tested. Describe any additional tests conducted, if applicable. Please refer the Model Testing Guidance for back-test requirements as per model usage and model forecast horizon. If the back-testing methodology is not fully specified by the Model Testing Guidance, provide justification as to why the back-testing scheme used is satisfactory for model performance testing.]

Yes, CAMU conducted backtesting on different time periods. The time periods cover a wide range of macroeconomic environment shifts, from the Financial Crisis of 2008-2009to the recent period marked with significant increases in housing prices and falling unemployment. Thus, backtesting sufficiently covers potential model risks within its performance and use.

Recent non-stress experience is captured through twelve (12) month analysis on the June 2017 portfolio, twenty-seven (27) month analysis on the March 2016 portfolio and the sixty (60) month analysis on the June 2013 portfolio snapshots. Performance under a stressed environment is captured through twelve (12), twenty-seven (27) month and sixty (60) month analyses on the January 2008 portfolio snapshot. (See below for discussion of time frames.) The 60 month backtest results for stress and non-stress also addressed the CECL requirements which requires backtest results over the life of the loan. Please refer to the backtest attachments mentioned in the narrative below for the 60 month backtest results. CAMU additionally conducted backtest analysis for the OOT period(snapshot 201201) for the 27 month horizon. Other backtest results, in compliance with MRM's template design have been separately uploaded on the iMRMS server.

- Describe the time frame of the back-testing and justification of the choice.

  [Describe the data used for back-testing. Explicitly describe the data, and explain the relevance to the model.]

  As per the model usage grid (Section 1.1) and the Model Testing Guidance, CAMU conducted backtesting over multiple time horizons:

    - Short (12 months)
    - Intermediate (27 months)
    - Long (60 months)

  The back-testing is performed on the entire population as of each tested portfolio snapshot. The macroeconomic data is time varying, and the behavioral data is static as of the snapshot month. The backtest has been conducted for both stress and most recent period, which are defined as follows-

  - Most Recent: to test performance of the model under recent conditions
  - Stress: to test performance of the model under stressed conditions, as defined by Gating Principles

The model predictions are compared to actual values to assess the model accuracy for each of these backtest horizon (Predicted numbers extracted from a sub-period of a long-term run may be slightly different from predicted numbers directly from a specific short period run due to loss timing distribution.)

Table 6.3.1.1: Backtesting Performance Horizons

Back test for RM(Non-MOD)	Snapshot date	first forecast month	Horizon (months)
Stress back test	200801	200802	1. 12 2. 27 3. 60
OOT back test	201201	201202	1. 27
Recent 60 month back test	201306	201307	1. 60
Recent 27 month back test	201603	201604	1. 27
Recent 12 month back test	201706	201707	1. 12

- Has back-testing been conducted on a granular level - at the logical segments/product level?

  [Explain what segmentation was considered for back-testing. Highlight all low-default/immaterial segments as per the Model Testing Guidance, and explain how these segments were treated in back-testing performance and interpretation of back-testing results. If certain segments were omitted, provide the rationale and demonstrate that these segments do not have a material impact.]

Yes, CAMU conducted backtesting at the aggregate portfolio level and at logical segment levels that not only validates and aligns with the model's segmentation scheme but moves beyond the chosen segmentation structure to incorporate additional risk drivers that are considered relevant to the business. For the 2019 CCAR process, CAMU conducted characteristic analysis according to the following segments and risk drivers-

- Business Entity (CPB vs. CMI)
- Refreshed FICO
- Combined Loan-to-Value
- Delinquency
- Region
- ARM vs Fixed Loan
- Corp vs Holdings

For additional details, please refer to Section 6.3.2 of the MDD.

It should be noted that, given the rich loan-level performance history, there could be infinite ways to dissect the portfolio and analyze its behavior at granular levels. Excessive granularity can stretch the data too thin and lead to unintuitive inferences. CAMU has strived to maintain a balance in its granular level reporting structure by adhering to those risk drivers that-

1. Significantly affect portfolio performance
2. Recommended by model end-users and model reviewers'(external + internal)
3. Based on industry wide trends and other literature review
4. Align with business intuition and are statistically significant

- Have the recommended back-testing metrics been used? If thresholds are breached, provide rationale for why and ensure that the model risk is mitigated by model developer. If any of these tests are not applicable, please explain why and please provide alternative analyses / tests to overcome these challenges.

  [Describe the metrics used if they differ from the Model Testing Guidance. Explain evaluation of prediction error with respect to the model components for any of the tests not provided in the Model Testing Guidance.]

  Yes, the recommended back-testing metrics have been used for evaluating the model's performances across different snapshots. The Cumulative Percent Prediction Error (CCERPCT) is the main parameter to measure model accuracy. The CCERPCT Error threshold differs across the different time horizons over which backtesting were conducted. If the Cumulative percent prediction error for the short term, medium term and long term horizons were below 20%, 25% and 40%, respectively, the model is considered as acceptable and no special actions are required.

Table 6.3.1.2: Backtesting Performance Metric Definitions

Performance Measure	Periodic	Cumulative
Actual Outcome	$A(i)$	$CA(i) = \sum_{j=1}^{n} A(j)$
Predicted Outcome	$P(i)$	$CP(i) = \sum_{j=1}^{n} P(j)$
Prediction Error	$ERR(i) = P(i) - A(i)$	$CERR(i) = CP(i) - CA(i)$
Percent Prediction Error	$ERRPCT(i) = \frac{ERR(i)}{A(i)}$	$CERRPCT(i) = \frac{CERR(i)}{CA(i)}$
Mean Absolute Deviation	$MAD = \frac{1}{N}\sum_{i=1}^{N} abs[ERR(i)]$	
Root Mean Square Error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} ERR(i)^2}$	
Coefficient of Variation	$COV = \frac{RMSE}{\frac{1}{N}CA(N)}$	

Performance Measure	Periodic	Cumulative

6.3.2 Back-testing results

[Provide a summary of key model performance metrics, which can be found in the Model Testing Guidance. Comment on model performance, including its pros and cons in stress and non-stress periods. Provide business justification of model performance – link model performance to changes in portfolio structure and/or portfolio behavior. Comment on performance of key segments.]

[Explicitly state whether the back-testing results are within acceptable thresholds as defined in the Model Testing Guidance.]

[Explain breaches of back-testing thresholds and provide business justification, and describe subsequent limitations to the model usage and mitigation actions for model use in production.]

The end-to-end RM model suite performs reasonably well on both unit and dollar predictions for delinquency buckets, Voluntary Prepayment (VP), Exposure at Default (EAD), Default (Gross Credit Loss – GCL), and Net Credit Loss (NCL).

Before delving deeper into the 2019 backtest results, it is important to provide some historical context around the prior year's model results and the subsequent improvements that were targeted this modeling cycle to address the prior year model's feedback.

The prior version of the Method A first lien model was robust and provided reasonably strong performance across all portfolio snapshots. For a range of portfolio snapshots and market environments, the model accurately predicted the most important metrics of NCL and GCL, as was measured in dollar terms. Consequently, based on the review of the model results by the relevant stakeholders' certain limitations (#s 16575, 16577) were raised as the model was found to under predict results for some final forecasted variables and/or individual transition equations, across the non-stress and/or stress period backtest for the short, medium and/or long term forecast horizons.

Given the feedback on the underestimation of model results across stress, (the prior models had underestimated stress period results for three consecutive years) and non-stress, the 2019 modeling efforts were specifically targeted to address the underestimation and improve the model's macroeconomic sensitivity across both stress and recent timeframes.

As previously discussed in Sections 1.1 and 5.2, the following enhancements were attempted on the 2019 RM model to specifically improve the model's estimation and sensitivity in both stress and recent periods-

1. The development sample includes the most recent performance history until end of 2017. The inclusion of recent performance data has contributed significantly to model performance improvement especially given the significant loan composition and performance changes in the recent years compared to historical stress.

2. A methodical approach was utilized to segregate between the modeled vs non-modeled transitions. Because of adopting this new approach, some of the off-diagonal transitions were now modeled and additionally separate logics were introduced to systematically account for the stress vs recent period for segments of interest (conventional vs government, in trial vs not in trial) differentials. Please refer to Section 5.1.2 for additional details.

3. The segmentation scheme was completely re-developed using quantitative factual based analytical tools (decision tree) which led to the introduction of several new interaction variables that would be able to capture the model's recent and stress periods performance better across granular segments. Please refer to Section 5.1.3 for additional details.

4. The prior year's DV model has been replaced with a new DV haircut lookup logic. The new DV logic is more sensitive to macro-economic factors compared to last year's model. Please refer to MDD Sections 3.2 and 5.1.2 for additional detail.

The key motivation behind these definitive improvements was to obtain better prediction in recent period and improve the model's stress period accuracy.

The model summary results are presented in Table 6.3.2.1 below.

Table 6.3.2.1: Summary of Back-testing Results

Model component/ segment	Type of test	Test results (breached parameter)	Business justification/ Limitations	Mitigation*
CMI - Recent & Stress	Full Sample (12/27/60 months)	GCL/NCL ($MM)	One Threshold Breach in NCL(12month) with small dollar impact	Overlay if necessary
CPB - Recent & Stress	Full Sample (12/27/60 months)	GCL/NCL ($MM)	Few Breaches across Snapshots, loss amount is immaterial, limited data, conservative estimate	Overlay if necessary
Overall Model (CMI/USRB/CPB)	OOT Sample (27 months)	GCL/NCL ($MM); GCL units; EAD Balance	No Threshold Breaches	Not applicable.
EAD	Full Sample (12/27/60 months)	Balances	No Threshold Breaches	Not applicable.
Intermediate Delinquencies , IVP, VP	Full Sample (12/27/60 months)	Units and Balances	No Threshold Breaches	Not applicable.
SFO – Recent & Stress	Full Sample (12/27/60 months)	GCL($MM)	Few Breaches across Snapshots, conservative estimate	Overlay if necessary

Details on MRM server locations for model runs and summary reports used for numbers presented in the MDD are provided in the excel file with name starting '4.1.9_Production Upload Listing.xlsx'.

- Provide detailed back-test results.

[Provide detailed results of all the back-tests performed, and provide a brief commentary on model performance. All the backtest results should be provided in backtest excel template as embedded in the Model Testing Guidance]

This sub-section below contains the full sample back-test results for both the stress and recent periods for all the three-backtesting horizons (short, medium and long) for GCL and NCL metrics. Both of these metrics assess the model's component level and end-to-end performance. Results have been summarized using tables and figures as shown below. Also presented are the EAD, LGD component model testing and Characteristics analyses.

- Do quarterly back-test results show evidence of error trends over time, or high quarterly volatility? If so, what remediation is planned. Please refer to model testing guidance for more details.

CAMU reviews quarterly model performance monitoring results with stakeholders from model sponsor teams (including model end users). CAMU and model sponsor teams deem that monitoring the respective portfolios and utilizing model outputs along with overlays approved by MRM would be sufficient compensating controls for any breaches.

6.3.2.1 Full Sample Backtest Results

Presented below within Figures 6.3.2.1.1 – 6.3.2.1.4 are the model backtest results at CMI and CPB segment levels for GCL and NCL metrics. Please note the GCL numbers capture the PD model performance while NCL captures the end-to-end model performance.

Table 6.3.2.1.1: Backtesting Results –GCL: Recent and Stress Scenarios (CMI)

Backtest Period	Model Usage	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold	CERRPCT Result	Final Decision
Stress 200801 12 months	Short Term Uses	7,098.43	7,003.40	(95.03)	-1.34%	20%	Pass	Back-testing Pass
Stress 200801 27 months	Medium Term Uses	18,905.65	17,902.63	(1,003.03)	-5.31%	25%	Pass	Back-testing Pass
Stress 200801 60 months	Long Term Uses	27,049.77	25,364.54	(1,685.23)	-6.23%	40%	Pass	Back-testing Pass
Recent 201706 12 months	Short Term Uses	373.59	378.66	5.07	1.36%	20%	Pass	Back-testing Pass
Recent 201603 27 months	Medium Term Uses	1,023.53	1,066.98	43.45	4.24%	25%	Pass	Back-testing Pass
Recent 201306 60 months	Long Term Uses	3,239.33	3,598.73	359.40	11.09%	40%	Pass	Back-testing Pass

Table 6.3.2.1.2: Backtesting Results –NCL: Recent and Stress Scenarios (CMI)

Backtest Period	Model Usage	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold	CERRPCT Result	Final Decision
Stress 200801 12 months	Short Term Uses	2,183.65	2,237.37	53.72	2.46%	20%	Pass	Back-testing Pass
Stress 200801 27 months	Medium Term Uses	5,942.94	6,146.32	203.38	3.42%	25%	Pass	Back-testing Pass
Stress 200801 60 months	Long Term Uses	8,876.20	8,890.82	14.62	0.16%	40%	Pass	Back-testing Pass
Recent 201706 12 months	Short Term Uses	14.45	10.97	(3.48)	-24.08%	20%	Fail	Back-testing Fail
Recent 201603 27 months	Medium Term Uses	49.90	53.27	3.37	6.76%	25%	Pass	Back-testing Pass
Recent 201306 60 months	Long Term Uses	320.22	361.13	40.91	12.78%	40%	Pass	Back-testing Pass

Table 6.3.2.1.3: Backtesting Results –GCL: Recent and Stress Scenarios (CPB)

Backtest Period	Model Usage	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold	CERRPCT Result	Final Decision
Stress 200801 12 months	Short Term Uses	17.25	106.11	88.86	515.21%	20%	Fail	Back-testing Fail
Stress 200801 27 months	Medium Term Uses	225.82	359.58	133.76	59.23%	25%	Fail	Back-testing Fail
Stress 200801 60 months	Long Term Uses	479.67	612.59	132.93	27.71%	40%	Pass	Back-testing Pass
Recent 201706 12 months	Short Term Uses	78.54	54.04	(24.50)	-31.19%	20%	Fail	Back-testing Fail
Recent 201603 27 months	Medium Term Uses	118.06	106.54	(11.53)	-9.76%	25%	Pass	Back-testing Pass
Recent 201306 60 months	Long Term Uses	244.46	264.44	19.98	8.17%	40%	Pass	Back-testing Pass

Table 6.3.2.1.4: Backtesting Results –NCL: Recent and Stress Scenarios (CPB)

Backtest Period	Model Usage	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold	CERRPCT Result	Final Decision
Stress 200801 12 months	Short Term Uses	0.34	14.74	14.41	4298.78%	20%	Fail	Back-testing Fail
Stress 200801 27 months	Medium Term Uses	6.40	63.18	56.77	886.47%	25%	Fail	Back-testing Fail
Stress 200801 60 months	Long Term Uses	52.85	115.48	62.63	118.50%	40%	Fail	Back-testing Fail
Recent 201706 12 months	Short Term Uses	1.40	3.97	2.58	184.00%	20%	Fail	Back-testing Fail
Recent 201603 27 months	Medium Term Uses	5.38	6.54	1.16	21.60%	25%	Pass	Back-testing Pass
Recent 201306 60 months	Long Term Uses	23.83	25.28	1.44	6.06%	40%	Pass	Back-testing Pass

Stress Period Performance

For the CMI RM portfolio (Figures 6.3.2.1.1 and 6.3.2.1.2), the model generates reasonably accurate forecasts for GCL estimates across all testing horizons, all of which are within the stipulated thresholds. The forecasts for stress NCL are on the conservative side due to the targeted model improvements made this year to specifically address the prior model's stress period under-estimation. Please note that these estimates did not include the $350MM FFIEC one-time bulk entry adjustment related to non-performing asset sale transactions, which occurred in March 2010. This one-time adjustment is not at loan level, and has been acknowledged and verified by both Finance and Loss Forecasting teams. Therefore, the error percentage of medium and long term back testing results on the January 2008 snapshot would be reduced by $350MM to reflect this overall level NCL adjustment (adding $350MM) to actual losses.

For the CPB RM portfolio (Figures 6.3.2.1.3 and 6.3.2.1.4), the model continues to produce conservative stress period forecast due to limited data availability on actual losses. As well discussed in earlier sections of the MDD, the CPB portfolio has always exhibited extremely low and volatile loss volume over time. To add some context, there were only 254 first lien CPB loans with losses during the entire history of 2008-2017. Further, the lack of CPB losses during the historical stress period (in particular, there were only 20 residential mortgage loans with $9MM lifetime losses during the recession 2008-2009) implied that there can be no statistical model that can yield reasonable predictions for such a sparse volume of loans. As such, the model continues to over-estimate the CPB stress losses.

Recent Period Performance

For the CMI RM portfolio alone, the model stays within threshold limits, producing slightly conservative estimates for GCL across all testing horizons. For NCL, the model predicts reasonably well for medium term and long-term horizons. The short-term breach in NCL has relatively low dollar impact ($4MM). The four-quarter tracking period, together with relatively low actual loss level has posed challenges to accurately forecast both loss amount and loss timing. Further analysis also shows that the under-estimation comes from the Holding portfolio whereas the model performs fairly well on the Corp portfolio. The Holdings portfolio is the de facto carrier of non-performing, riskier (sub-prime) originations, which are tactfully managed for their incremental riskiness and are phased out for gradual future asset sales, depending on market timing and competition.

For the CPB RM portfolio, given the cited limitation on CPB's significant overestimation in prior year's model for recent periods, diligent efforts were made this modeling cycle to introduce specific macro-economic and balance effects to improve the current model's forecast. While the current model successfully reduced the significant over-estimation of recent periods as witnessed in prior year's model, it produces a 22% over-estimation for NCL in the medium term. Although the relative impact of the error breach seems high, the absolute dollar amount reflecting this breach is approx $1.2MM. In addition, it is important to note that the NCL values (absolute ~$5.38MM and relative to portfolio size) do not meet the materiality thresholds (NCL ~ $10MM per Model Testing Guidance). The nature of the CPB portfolio – credit quality, limited size) leads to sparse IVP and loss experience such that a couple of loans can dramatically affect the actual-versus-predicted comparisons in backtesting.

Back-Testing Summary for PD Component Model – Intermediate Delinquencies, IVP, VP (Units and Balances)

Based on the model's Usage Grid, some of the non-CCAR usages of the model correspond to leveraging the transition model's monthly delinquencies, IVP and VP balances and units for BAU reporting and loss forecasting requirements. Hence, CAMU evaluated the PD's model accuracy in forecasting the delinquency statuses, IVP and VP transitions. Tables 6.3.2.1.5a and 6.3.2.1.5b summarizes the CERRPCT for both balances and units in each loan state over the twenty-seven (27) month horizon beginning in March 2016 for the CMI and CPB RM portfolios respectively. The table reports the error of GCL (IVP balance plus 27-month averaged bucket7 balance), terminated outcomes of default (IVP) and prepays (VP), as well as all-intermediate delinquent buckets. Overall, as can be noted the model performs reasonable well across all transitions.

Table 6.3.2.1.5a: Backtesting Results – Intermediate Delinquencies, IVP, VP (Units and Balances) – CMI Only

NA Mortgage Method A - Cummulative Error % Report
CMIRM_Combined

	Recent 201706 12 Months (Threshold 20%)	Recent 201603 27 Months (Threshold 25%)	Recent 201306 60 Months (Threshold 40%)	Stress 200801 12 Months (Threshold 20%)	Stress 200801 27 Months (Threshold 25%)	Stress 200801 60 Months (Threshold 40%)
BAL BK01 CERRPCT	-1.07%	0.28%	2.82%	0.00%	1.80%	3.86%
Unit BK01 CERRPCT	-0.75%	-0.91%	-3.44%	-0.84%	-0.65%	-1.41%
BAL BK2 CERRPCT	-4.40%	5.04%	3.25%	-2.50%	-7.82%	-10.13%
Unit BK2 CERRPCT	1.42%	8.16%	0.86%	-4.51%	-7.31%	-10.10%
BAL BK3 CERRPCT	-6.00%	9.34%	13.47%	-4.17%	-9.81%	-10.59%
Unit BK3 CERRPCT	-10.30%	1.95%	4.09%	-7.42%	-7.60%	-11.01%
BAL BK4 CERRPCT	-13.20%	5.61%	28.34%	-5.06%	-17.51%	-18.81%
Unit BK4 CERRPCT	-18.03%	-1.81%	14.05%	-8.13%	-18.11%	-18.47%
BAL BK5 CERRPCT	-12.80%	5.93%	25.76%	-3.13%	-16.20%	-16.64%
Unit BK5 CERRPCT	-13.25%	-1.21%	12.40%	-5.24%	-12.75%	-16.00%
BAL BK6 CERRPCT	-8.32%	12.38%	29.36%	-5.17%	-18.25%	-16.09%
Unit BK6 CERRPCT	-10.37%	0.97%	12.88%	-6.35%	-14.22%	-15.01%
BAL BK7 CERRPCT	5.12%	16.26%	21.20%	11.13%	-1.12%	20.08%
Unit BK7 CERRPCT	2.07%	8.71%	10.96%	19.10%	4.82%	14.77%
BAL IVP CERRPCT	-2.50%	1.58%	20.67%	-15.01%	14.01%	35.78%
Unit IVP CERRPCT	-2.32%	2.18%	18.81%	-5.45%	16.07%	26.08%
BAL VP CERRPCT	28.32%	8.64%	-0.04%	15.58%	4.25%	3.01%
Unit VP CERRPCT	11.30%	5.54%	8.03%	16.69%	7.37%	8.33%
Unit GCL CERRPCT	-2.76%	-0.58%	8.00%	4.37%	1.04%	-1.54%
acl_gcl	$374M	$1,024M	$3,239M	$7,098M	$18,906M	$27,050M
pred_gcl	$379M	$1,067M	$3,599M	$7,003M	$17,903M	$25,365M
GCL CERRPCT	1.36%	4.24%	11.09%	-1.34%	-5.31%	-6.23%
Actual_EAD	$46,575M	$47,600M	$45,072M	$130,272M	$126,339M	$114,660M
Fckz_EAD	$46,860M	$48,206M	$46,605M	$130,770M	$127,982M	$120,283M
EAD CERRPCT	0.61%	1.29%	3.40%	0.38%	1.31%	4.90%
ACT_NCL	$16M	$50M	$320M	$2,184M	$5,943M	$8,876M
PRED_NCL	$11M	$53M	$361M	$2,237M	$6,146M	$8,891M
NCL CERRPCT	-34.08%	6.76%	12.78%	2.46%	3.42%	0.16%

Business / Model / Segment
CMIRM_Combined

BT_Period
All

Term
All

snapdate
All

Measure Names
Multiple values

Table 6.3.2.1.5b: Backtesting Results – Intermediate Delinquencies, IVP, VP (Units and Balances) – CPB

NA Mortgage Method A - Cummulative Error % Report
CPB_RM

	Recent 201706 12 Months (Threshold 20%)	Recent 201603 27 Months (Threshold 25%)	Recent 201306 60 Months (Threshold 40%)	Stress 200801 12 Months (Threshold 20%)	Stress 200801 27 Months (Threshold 25%)	Stress 200801 60 Months (Threshold 40%)
BAL BK01 CERRPCT	-0.70%	0.40%	0.30%	-1.52%	-0.47%	12.68%
Unit BK01 CERRPCT	-0.87%	-0.56%	-0.84%	-2.97%	-2.42%	10.47%
BAL BK2 CERRPCT	32.78%	31.02%	16.22%	10.94%	37.15%	28.37%
Unit BK2 CERRPCT	40.32%	41.80%	23.65%	10.29%	27.60%	26.19%
BAL BK3 CERRPCT	18.28%	2.55%	18.17%	50.37%	47.51%	24.38%
Unit BK3 CERRPCT	9.40%	6.87%	18.35%	52.25%	47.21%	20.69%
BAL BK4 CERRPCT	27.87%	12.70%	17.74%	71.27%	50.53%	23.84%
Unit BK4 CERRPCT	13.13%	0.73%	0.77%	85.90%	65.99%	26.07%
BAL BK5 CERRPCT	-10.59%	-4.66%	12.19%	107.91%	55.33%	21.15%
Unit BK5 CERRPCT	-7.66%	-4.11%	6.82%	152.65%	97.23%	40.38%
BAL BK6 CERRPCT	-22.42%	-6.05%	17.57%	291.34%	56.99%	22.92%
Unit BK6 CERRPCT	-9.66%	7.04%	17.50%	219.47%	102.42%	43.61%
BAL BK7 CERRPCT	-10.27%	9.80%	36.00%	285.00%	57.64%	31.78%
Unit BK7 CERRPCT	-5.65%	22.62%	38.40%	202.50%	134.30%	64.04%
BAL IVP CERRPCT	6.30%	-4.45%	0.30%	380.03%	222.38%	166.92%
Unit IVP CERRPCT	56.07%	-6.90%	-2.15%	1122.35%	394.76%	170.04%
BAL VP CERRPCT	25.50%	11.05%	4.51%	47.57%	2.00%	-22.08%
Unit VP CERRPCT	7.99%	0.88%	-4.28%	45.57%	-10.54%	-32.24%
Unit GCL CERRPCT	-4.56%	-0.13%	12.08%	455.05%	132.76%	64.23%
act_gcl	$79M	$118M	$244M	$17M	$226M	$480M
pred_gcl	$54M	$107M	$264M	$106M	$360M	$613M
GCL CERRPCT	-31.19%	-9.76%	8.17%	515.21%	59.23%	27.71%
Actual_EAD	$23,460M	$19,728M	$14,450M	$16,710M	$16,312M	$15,621M
Fcbz_EAD	$23,689M	$20,079M	$14,919M	$16,923M	$16,772M	$16,434M
EAD CERRPCT	0.04%	1.78%	3.25%	1.28%	2.82%	5.21%
ACT_NCL	$1M	$5M	$24M	$0M	$6M	$53M
PRED_NCL	$4M	$7M	$25M	$15M	$63M	$115M
NCL CERRPCT	184.00%	21.60%	6.06%	4298.78%	886.47%	118.50%

Business / Model / Segment
CPB_RM

BT_Period
All

Term
All

snapdate
All

Measure Names
Multiple values

Back-Testing – EAD

Please refer to Tables 6.3.2.1.5a and 6.3.2.1.5b for EAD backtest results across all tested horizons. All EAD CERRPCT errors are very low and within MRM's prescribed horizon threshold limits.

6.3.2.2 Back-Testing Summary for LGD Component Model

Presented below are the LGD component level backtesting results. The backtesting has been conducted on a consistent set of snapshots for the short, medium and long-term horizons across first liens, second liens and the combined portfolio. The component level LGD model shows reasonable backtesting performance across all testing horizons (both recent and stress periods) without any threshold breaches. Please refer to the following attachments for additional details - '6.3.2.1 Severity Model_Backtest_Detailed.xlsx' and '6.3.2.1 Severity Model_Backtest_Summary.xlsx'.

Recent Period Backtest

Table 6.3.2.2.1: LGD 12 Month Back-testing Results - June 2017 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	12 Month Backtest	Within 20% 18.31%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	12 Month Backtest	Within 20% 3.91%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	12 Month Backtest	Within 20% 7.70%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

Table 6.3.2.2.2: LGD 27 Month Back-testing Results – March 2016 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	27 Month Backtest	Within 25% 8.60%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	27 Month Backtest	Within 25% 1.59%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	27 Month Backtest	Within 25% 3.44%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

Table 6.3.2.2.3: LGD 60 Month Back-testing Results – June 2013 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	60 Month Backtest	Within 40% 13.25%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	60 Month Backtest	Within 40% -0.74%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	60 Month Backtest	Within 40% 3.50%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

Stress Period Backtest

Table 6.3.2.2.4: LGD 12 Month Back-testing Results - Jan 2008 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	12 Month Backtest	Within 20% 7.08%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	12 Month Backtest	Within 20% 0.81%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	12 Month Backtest	Within 20% 2.62%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

Table 6.3.2.2.5: LGD 27 Month Back-testing Results – Jan 2008 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	27 Month Backtest	Within 25% 6.36%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	27 Month Backtest	Within 25% 0.64%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	27 Month Backtest	Within 25% 2.22%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

Table 6.3.2.2.6: LGD 60 Month Back-testing Results – Jan 2008 Snapshot

Model component/ segment	Type of test	Test results CERRPCT	Business justification/ Limitations	Mitigation
First Lien	60 Month Backtest	Within 40% 5.96%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Second Lien	60 Month Backtest	Within 40% 0.20%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.
Combined	60 Month Backtest	Within 40% 1.95%	Not applicable. Since it is within model performance threshold.	Not applicable, since there is no limitation.

6.3.2.3 Characteristic Analysis Results

Characteristic analysis is a measurement of the model's performance (back-test accuracy) to key risk drivers. The results cover 27 month forecast horizons beginning in March 2016, referred to as "Recent," and January 2008, referred to as “Stress.” The characteristic analyses demonstrate that the model appropriately captures the sensitivity and risk rank ordering across key risk drivers and segments measured by both GCL rate and NCL rate. The impact of segments with relatively bigger errors is mitigated by the small portfolio allocation (# account share) and(or) the small loss shares. It is worth noting that the model performs reasonably well across Corp vs Holdings split in recent period as illustrated in Figures 6.3.2.3.1.The model risk-ranks the split accurately with Corp loans depicting better performance compared to Holdings. This conforms to business intuition as the Corp. Portfolio specifically houses loans with stronger credit quality. Please note there was no Corp/Holdings split at the time of crisis. This split was an aftereffect of the Historical crisis.

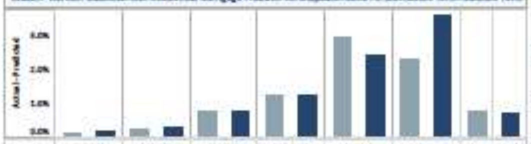
Please note Figure **6.3.2.3.1** has been presented in a GRID View format for recent period for the RM CMI Portfolio only for all other dimensions. The CMI Stress Period chart can be found in attachment – ‘6.3.2 Characteristic Analysis_CMI_Stress.pdf’. The CPB Characteristics analysis charts(recent and stress) are available in the following attachments – ‘ 6.3.2 Characteristic Analysis_CPB_Recent.pdf’ and ‘6.3.2 Characteristic Analysis_CPB_Stress.pdf’.

Figure 6.3.2.3.1: Characteristic Analysis – GCL/NCL Rates (Recent Period) – CMI Only

CLTV BKT

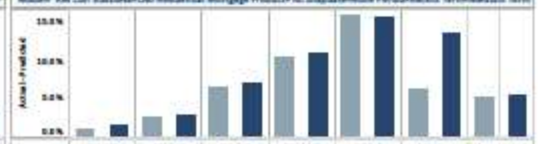
roll_name	curr_business	product	conventional	term	Period
RM	CMI-Residential Mortgage	All	All	Medium Term	Recent

NCL Actual, Predicted and Error by CLTV_BKT
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate=None Period=Recent Term=Medium Term



	(00-60]	(60-80]	(80-90]	(90-100]	(100+)	Missing	Aggregated
Act Loss Rate	0.10%	0.27%	0.77%	1.20%	2.87%	2.28%	0.79%
Pred Loss Rate	0.18%	0.30%	0.78%	1.22%	2.41%	3.60%	0.72%
Acct Share	44.32%	31.18%	10.14%	8.39%	15.89%	0.11%	100.00%
Loss Share	4.20%	16.79%	9.99%	13.13%	38.37%	0.31%	100.00%

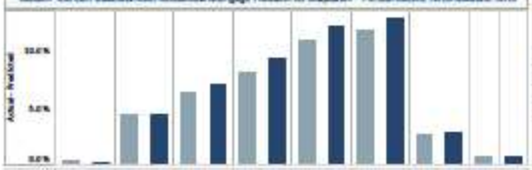
GCL Actual, Predicted and Error by CLTV_BKT
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate=None Period=Recent Term=Medium Term



	(00-60]	(60-80]	(80-90]	(90-100]	(100+)	Missing	Aggregated
Actual GCL Rate	1.12%	3.61%	6.48%	10.67%	15.82%	6.07%	5.27%
Predicted GCL Rate	1.40%	2.99%	7.00%	10.88%	15.50%	13.88%	5.58%
Acct Share	44.21%	31.13%	10.10%	8.33%	15.80%	0.12%	100.00%
Loss Share	3.30%	18.78%	9.45%	15.29%	50.07%	0.31%	100.00%

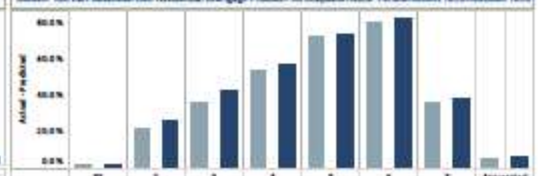
Mba Delq Status

NCL Actual, Predicted and Error by Mba_delq_Status
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate= * Period=Recent Term=Medium Term



	01	2	3	4	5	6	7	Aggregated
Act Loss Rate	0.03%	3.88%	6.01%	8.10%	11.00%	11.82%	2.71%	0.79%
Pred Loss Rate	0.29%	4.34%	7.05%	9.94%	12.23%	12.94%	2.78%	0.72%
Acct Share	96.48%	1.89%	0.81%	0.39%	0.69%	0.42%	0.31%	100.00%
Loss Share	46.18%	12.18%	7.50%	4.47%	6.79%	4.18%	18.03%	100.00%

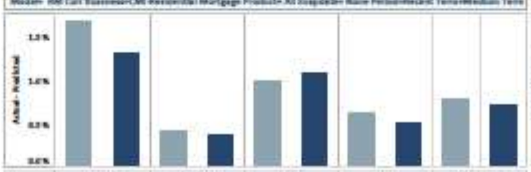
GCL Actual, Predicted and Error by MBA_delq_Status
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate=None Period=Recent Term=Medium Term



	01	2	3	4	5	6	7	Aggregated
Actual GCL Rate	1.02%	21.30%	36.07%	54.53%	72.70%	80.09%	36.09%	5.27%
Predicted GCL Rate	1.79%	23.82%	42.02%	56.97%	73.40%	82.37%	37.90%	5.20%
Acct Share	96.48%	1.89%	0.92%	0.39%	0.49%	0.51%	5.26%	100.00%
Loss Share	45.14%	10.18%	7.88%	4.37%	6.79%	4.03%	18.03%	100.00%

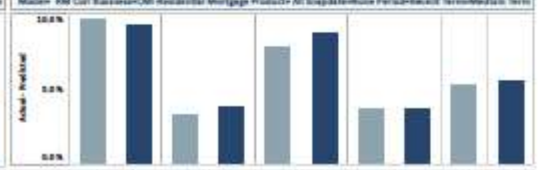
Prop_Region

NCL Actual, Predicted and Error by Prop_Region
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate= None Period=Recent Term=Medium Term



	Midwest	Northeast	South	West	Aggregated
Act Loss Rate	1.72%	0.41%	1.00%	0.66%	0.79%
Pred Loss Rate	1.31%	0.38%	1.08%	0.52%	0.72%
Acct Share	11.61%	24.79%	22.47%	39.88%	100.00%
Loss Share	24.91%	11.97%	28.98%	31.82%	100.00%

GCL Actual, Predicted and Error by Prop_Region
Model= RM Curr Business=CMI-Residential Mortgage Product= All Snapdate=None Period=Recent Term=Medium Term



	Midwest	Northeast	South	West	Aggregated
Actual GCL Rate	9.10%	3.28%	7.81%	3.79%	5.27%
Predicted GCL Rate	9.22%	3.77%	8.70%	3.40%	5.58%
Acct Share	11.61%	24.79%	22.97%	39.88%	100.00%
Loss Share	29.91%	11.97%	28.86%	31.81%	100.00%

Please note that the next four (4) charts – Figure **6.3.2.4.2** -**6.3.2.4.5** correspond to the split by business (CMI vs CPB). The model risk-ranks the portfolios (CMI vs CPB) accurately. Recent period CPB results for NCL are outside of threshold but are also outside of materiality thresholds, with actual losses over the forecast horizon less than $10MM. For Stress period, CPB results are not meaningful due to limited and volatile loss experience – small sample sizes.

Figure 6.3.2.4.2: Characteristic Analysis – GCL Rate (Recent Period) – Business Entity

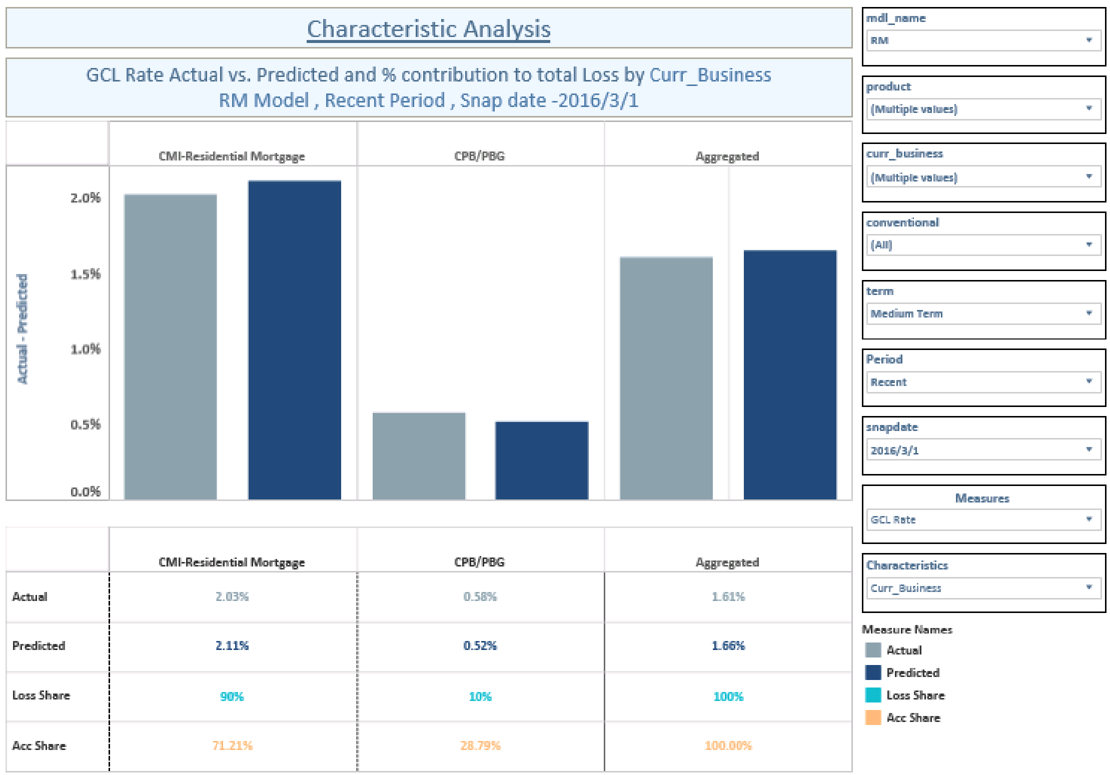


Figure 6.3.2.4.3: Characteristic Analysis – GCL Rate (Stress Period) – Business Entity

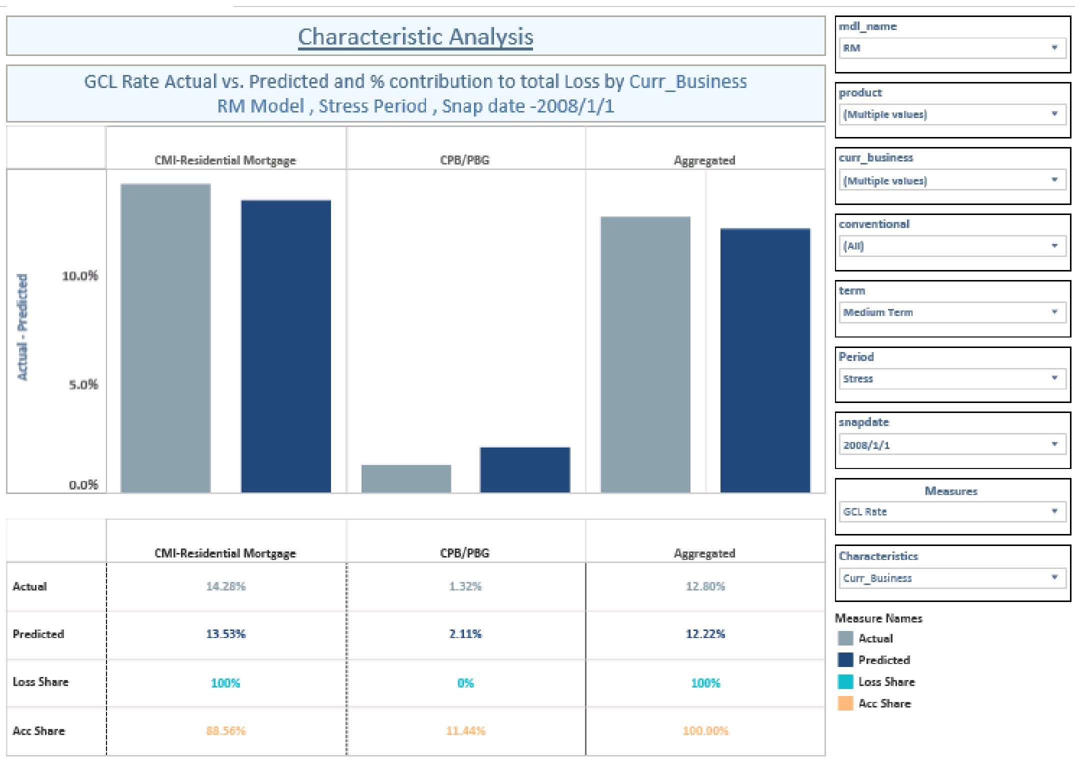


Figure 6.3.2.4.4: Characteristic Analysis – NCL Rate (Recent Period) – Business Entity

Characteristic Analysis

Loss (NCL) Rate Actual vs. Predicted and % contribution to total Loss by Curr_Business
RM Model , Recent Period , Snap date -2016/3/1

	CMI-Residential Mortgage	CPB/PBG	Aggregated
Actual	0.10%	0.03%	0.08%
Predicted	0.11%	0.03%	0.08%
Loss Share	90%	10%	100%
Acc Share	71.21%	28.79%	100.00%

mdl_name: RM
product: (Multiple values)
curr_business: (Multiple values)
conventional: (All)
term: Medium Term
Period: Recent
snapdate: 2016/3/1
Measures: Loss (NCL) Rate
Characteristics: Curr_Business

Measure Names
- Actual
- Predicted
- Loss Share
- Acc Share

Figure 6.3.2.4.5: Characteristic Analysis – NCL Rate (Stress Period) – Business Entity

Characteristic Analysis

Loss (NCL) Rate Actual vs. Predicted and % contribution to total Loss by Curr_Business
RM Model , Stress Period , Snap date -2008/1/1

	CMI-Residential Mortgage	CPB/PBG	Aggregated
Actual	4.49%	0.04%	3.98%
Predicted	4.64%	0.37%	4.16%
Loss Share	100%	0%	100%
Acc Share	88.56%	11.44%	100.00%

mdl_name: RM
product: (Multiple values)
curr_business: (Multiple values)
conventional: (All)
term: Medium Term
Period: Stress
snapdate: 2008/1/1
Measures: Loss (NCL) Rate
Characteristics: Curr_Business

Measure Names
- Actual
- Predicted
- Loss Share
- Acc Share

6.3.2.4 Back-testing results –SFO

As stated in section 1.2 of the MDD (Borrowed Model Approach), the Method A RM model has been borrowed for use on the SFO portfolio. Therefore, SFO back tests were conducted on both recent and stress snapshots using Method A RM model, in order to evaluate the model performance for the proposed SFO usage(s).

Table 6.3.2.5.1 below outlines the SFO portfolio back test design. All snapshots stay consistent with the RM Risk portfolio back testing. It is important to note here that the reportable metrics are tailored based on SFO business user inputs, mainly focused on the back test outcomes of the PD component.

Table 6.3.2.5.1: SFO Back Test Design Summary – Snapshot / Period and Key Metrics

Back test for SFO	Snapshot date	first forecast month	Horizon (months)
Stress back test	200801	200802	1. 12 2. 27 3. 60
Recent 60 month back test	201306	201307	1. 60
Recent 27 month back test	201603	201604	1. 27
Recent 12 month back test	201706	201707	1. 12

Table 6.3.2.5.2 below provides the GCL (AMT) for SFO recent and stress snapshots. As shown in Table 6.3.2.5.2, the model errors are all within MRM's error threshold during recent period. The threshold breaching for short term(12 month) stress period back test reflects the conservative stress forecast for that period.

Table 6.3.2.5.2: SFO Back Test Result Summary – GCL(AMT)

Backtest Period	Model Usage	Actuals ($MM)	Predicted ($MM)	CERR ($MM)	CERRPCT	CERRPCT Threshold	CERRPCT Result	Final Decision
Stress 200801 12 months	Short Term Uses	9,214.49	13,573.04	4,358.56	47.30%	20%	Fail	Back-testing Fail
Stress 200801 27 months	Medium Term Uses	33,267.56	40,117.93	6,850.37	20.59%	25%	Pass	Back-testing Pass
Stress 200801 60 months	Long Term Uses	54,443.07	65,300.20	10,857.13	19.94%	40%	Pass	Back-testing Pass
Recent 201706 12 months	Short Term Uses	503.35	447.18	(56.17)	-11.16%	20%	Pass	Back-testing Pass
Recent 201603 27 months	Medium Term Uses	1,709.89	1,590.74	(119.14)	-6.97%	25%	Pass	Back-testing Pass
Recent 201306 60 months	Long Term Uses	8,446.46	9,360.84	914.38	10.83%	40%	Pass	Back-testing Pass

6.3.2.5 Out of Time(OOT) Backtest Results

Presented in Table 6.3.2.5.1 are the OOT backtest results for the overall MOD portfolio. The OOT period is defined to be between Jan 2012 and Mar 2014. The rationale for selection this out of time sample has been thoroughly discussed in Section 4.1.3 of the MDD. The model demonstrates robust results across various back testing periods (stress, recent, OOT) over varied horizons (short, medium and long). It is worth noting that the OOT period results (shown above) are also well within MRM's error threshold for all relevant metrics. It should be noted that the OOT results are based on parameters that were finalized based on the in-sample and in-time model development data as discussed in Section 4.1.3. It is also a true blind back test as none of the parameters or assumptions were specifically calibrated based on the OOT period as the non-modeled transition values used the recent period non-modeled transition assumptions instead of assumptions customized specifically for the OOT period performance. The recent period non-modeled transition assumptions factor in the strong macro-economic environment in recent times with continuous home price appreciation, declining unemployment, much smaller portfolio size(CMI- Holdings). Further the strategic shift in the business which involved purging sub-prime quality loans from its balance sheet through periodic asset sales along with stringent underwriting criteria for new originations post the crisis greatly improved the portfolio composition mix; which is also likely to understate the risks associated with the 2012-2013 period, particularly for the transitions heavily driven by portfolio/risk management policies - such as the current or low DLQ to IVP transitions. Please note that despite this "blind" setting, the model still demonstrates satisfactory results during the OOT period, which serves as a testament to the model's parameter stability and robust performance over time. Please refer to attachment 6.3 Model OOT Backtests.xlsx for additional details and backtest performance on intermediate delinquency buckets, IVP,VP(balances and units included).

Table 6.3.2.5.1: Backtesting Results – OOT

Portfolio	Backtest Period	Tested Metric	Actuals	Predicted	CERR	CERRPCT	CERRPCT Result	Final Decision
Overall	OOT Period - 27 months backtest	NCL Balance ($MM)	$930	$852	($78)	-8.42%	Pass	Back-testing Pass
		GCL Balance ($MM)	$5,403	$5,689	$287	5.31%	Pass	Back-testing Pass
		GCL Units(#)	36,353	39,056	$2,703	7.44%	Pass	Back-testing Pass
		EAD($MM)	$1,925,933	$1,965,067	$39,135	2.03%	Pass	Back-testing Pass

6.4 Model Sensitivity

[The following summarizes the mandated Stress sensitivity testing as per model usage:

If applicable, the model's performance under well-defined, regulatory mandated or internally mandated stress scenarios should be provided. Using a scaled, simultaneous shock to the appropriate input parameters or key input variables to produce model output profile is an effective way to demonstrate reasonable model performance under the prescribed scenario.

The sensitivity results must be provided for the BHC as well as FRB scenarios. The Model Testing

Guidance has been embedded in section 8. It provides the principles and applicable test as per model usage.

Highlight model limitations as a result of the sensitivity testing. All the sensitivity testing results should be provided in the Sensitivity Template which is embedded in Model Testing Guidance.

The Method A Residential Mortgage Model demonstrates strong sensitivity to macroeconomic conditions. Borrower, collateral, and loan behavior respond to economic conditions. As suggested by intuition, borrower behavior (PD) is sensitive to financial conditions: interest rates, unemployment, and home prices. Additionally, LGD is directly impacted by home values. Sensitivity testing of Method A modeling parameters and key affirms such intuition.

For the 2019 Method A RM model, sensitivity analysis has been conducted using the June 2018 snapshot. Sensitivity testing has been conducted for the short, medium and long term testing horizons as shown below.

Snap date	first forecast month	Base / stress + Back test / sensitivity	Horizon (months)
201806	201807	Sensitivity testing runs	1. 12 2. 27 3. 60

6.4.1 Sensitivity of model results under baseline vs stress macro scenarios

[Provide evidence of overall sensitivity of the model dependent variable (for e.g. PD) and the final forecasted variable (for e.g. NCL, balance, revenue) projections under the stress scenario related to baseline scenario in terms of incremental of separation of PD/Losses/Balance/Revenue etc. The sensitivity results should be provided as per the model usage grid. Explain why forecasts are/are not sufficiently sensitive. See Model Testing Guidance for details]

The Method A RM Model demonstrates strong sensitivity to the macroeconomic environment. Stressed conditions (e.g., declining home prices, lower GDP, lower income, lower stock market index, higher unemployment) lead to higher loss expectations. Declined home prices increase the value of the borrower’s default option and depress the liquidation value of the property (collateral) in a default event. Increasing home prices implies increased equity in home thereby lowering incidence of default. A decrease in GDP reflects a slowdown in economic growth and decrease in per capita income/purchasing power, which increases the default incidence. Higher unemployment, lower income and lower stock market index increases the likelihood of stressed financial conditions for households and reduced borrower’s payment ability.

Have sensitivity tests been performed at various levels of the model - at the model component level (for e.g. PD, LGD, EAD), as well as final forecasted variable (for e.g. NCL/balance)?

[Describe which components were tested and justify if any of the components were omitted.]

Yes, CAMU conducted sensitivity tests on multiple levels.

GCL reflects default exposure over UPB, which essentially capture the PD model performance.

The LGD sensitivity analysis captures the LGD model's performance, without any additive effect of the PD model or EAD logic.

The EAD component analysis demonstrates the sensitivity of the EAD logic.

NCL, as the outcome variable affecting business performance, captures the totality of interactions and the end-to-end model suite's sensitivity, incorporating effects of PD, EAD logic, LGD respectively.

If the dependent variable and the final forecasted variable based on cumulative percent change between severely stressed and base scenario is less than threshold as per the Model Testing Guidance, provide justification and explain why you still consider the model to be sufficiently responsive to macro environment – is the path variation between severe stress and base scenario sufficient

[In case you find the path variation between severe stress and base scenario insufficient, run the sensitivity analysis on a more severe and carefully justified scenario in order to demonstrate that the model is sufficiently responsive to macro environment.]

Both (cumulative) GCL and NCL increased significantly in the Stress scenario versus the recent period; therefore, the model is sufficiently responsive to the macro environment.

If the dependent variable and the final forecasted variable based on cumulative percent change between severely stress and base scenario is less than threshold as per the Model Testing Guidance, provide justification and explain why you still consider the model to be sufficiently responsive to macro environment – is the model sufficiently responsive to macroeconomic stress?

[In case you find the model to be insufficiently responsive, analyze and remedy the model macro irresponsiveness through re-considering: 1) the macro factors being used in the model, 2) assumptions being made and 3) methods being used for model estimation.]

Both (cumulative) GCL and NCL increased significantly in the Stress scenario versus the Base scenario; therefore, the model is sufficiently responsive to the macro environment.

What are the sensitivity test results?

[Evaluate sensitivity metrics as defined in the Model Testing Guidance for base and stress comparisons. All the results must be justified and their business interpretation provided in the form of the following summary table.]

The Method A RM Model demonstrates significant sensitivity to the macroeconomic environment across CMI and CPB segments. This can be attributed to two main reasons-

1. The actual losses have continued trending down in recent years. As a result, the model's base forecast has reduced compared to one year ago with improved portfolio composition, continued home price appreciation and unprecedentedly low unemployment rate in recent years.

2. The improved model performance has narrowed the model errors in both stress and recent periods. The increased model accuracy has also contributed to the higher sensitivity ratio.

Presented below in Table 6.4.1.1 are the summarized sensitivity results. All additional details on the sensitivity test results across horizons and scenarios are available in the MRM Templates attachments that have been uploaded on the iMRMS server.

Table 4.4.1.1: Summary of Sensitivity Testing Results - Base vs. Stress

Model component	Test / shock description	Test results (metric applied and result)	Business interpretation	Mitigation / limitation if performance is not deemed sufficient
GCL	Base vs Stress Sensitivity Testing (27 month)	CMI and CPB	Adequately sensitive to Stress scenario	Not Applicable
NCL	Base vs Stress Sensitivity Testing (27 month)	CMI and CPB	Adequately sensitive to Stress scenario	Not Applicable
LGD component	Base vs Stress Sensitivity Testing (12 month)	18% - First Lien; 7% - Second Lien	Adequately sensitive to Stress Scenario for both liens	Not Applicable
LGD component	Base vs Stress Sensitivity Testing (27 month)	26% - First Lien; 8% - Second Lien	Adequately sensitive to Stress Scenario for both liens	Not Applicable
LGD component	Base vs Stress Sensitivity Testing (60 month)	25% - First Lien; 5% - Second Lien	Adequately sensitive to Stress Scenario for both liens	Not Applicable

Details on MRM server locations for model runs and summary reports used for numbers presented in the MDD are provided in the excel file with name starting '4.1.9_Production Upload Listing.xlsx'.

27-month Sensitivity Results

Presented below are the sensitivity test results for the 27 month testing horizon. As illustrated in Table 6.4.1.2, for the CMI RM portfolio, the stressed conditions represented by the Stress scenario, acting through the above-described dynamics, increase GCL by a factor of 2.7X and NCL by a factor of 13.8X. For the CPB portfolio, the GCL sensitivity is 3.3X and NCL is at 5.83X, the results presented in 6.4.2.3.

Table 6.4.1.2: Summary of Sensitivity Testing Results – Recent vs. Stress (CMI)

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	592.77	1.00	17.23	1.00
Stress	1,572.53	2.65	237.91	13.80

Table 6.4.1.3: Summary of Sensitivity Testing Results – Base vs. Stress (CPB)

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	117.21	1.00	15.32	1.00
Stress	386.61	3.30	89.37	5.83

Summary of LGD Component Sensitivity Results

For the LGD component, sensitivity analysis was conducted on the 201806 snapshot across the short, medium and long-term horizons across first and second liens. Please note lien position has been identified as one of the top nodes within the LGD Segmentation Analysis (Section 5.1.3). Summarized results are presented below. For detailed month-over month change, please refer to attachment-'Severity Model_Sensitivity.xlsx'. As shown in Table 6.4.2.4 below, the first lien portfolio exhibits greater sensitivity compared to second lien. For all testing horizons, second liens display less than 10% departure from baseline forecast. Looking at the 27-month horizon, the first lien segment is 3.2X more sensitive compared to second lien. The sensitivity results differences between first and second liens are not surprising as traditionally second liens tend to exhibit higher loss rates (under base), compared to first liens. As a result, the sensitivity to incremental losses under the stress scenario is much muted for second liens.

Table 5.4.1.4: Summary of Sensitivity Testing Results - Base vs. Stress – LGD Component

Average Loss ($) / quarter - 60Month	Scenario		% change from base
	Base	Stress	
First Lien	323,649	403,136	25%
Second Lien	235,606	248,255	5%
Average Loss ($) /	Scenario		% change

Average Loss ($) / quarter - 60Month	Scenario		% change from base
	Base	Stress	
quarter - 27Month	Base	Stress	from base
First Lien	327,705	411,688	26%
Second Lien	235,412	254,649	8%
Average Loss ($) / quarter - 12Month	Scenario		% change from base
	Base	Stress	
First Lien	329,818	387,650	18%
Second Lien	235,594	251,040	7%

6.4.2 Sensitivity of model results under stress macro scenarios vs historical stress episode

[Describe comparison of the model dependent variable and the final forecasted variable (for e.g. Loss/Balance/Revenue) projections under the hypothetical stress scenario(s) against observed experience under historical stress scenario. To the extent that there are significant differences in the level of the model dependent variable and the final forecasted variable (for e.g. Loss/Balance/Revenue) projections under hypothetical stress vs historically observed stress, provide explanation for whether this pattern is reasonable or not. For Example: Considerations may include different severity of historical vs. macro stress scenarios, changing portfolio composition over time, policies, etc. Additional runs to demonstrate such differences may include running a historical portfolio snapshot as of the current period vs. the current portfolio under the same scenario, running the current portfolio under both historical and stress scenarios, etc. Please refer to Model Testing Guidance for details on the calculations. Model Sponsor is required to provide sufficient rationale if the thresholds mentioned in the Model Testing Guidance are not met.]

In order to further assess the sufficiency of the model macroeconomic sensitivity, the forecasted NCL under the stress scenario were compared to the actual credit losses observed during the Great Recession of 2008-2009 for the CMI portfolio. The Method A RM portfolio results in the Stress scenario differs from the historical performance of the portfolio during the Great recession. This however is less a reflection of the model's performance but more a reflection of the considerable changes to the portfolio's composition

As illustrated in Figure 6.4.2.1, modeled loss rates under the Stress scenario are far below than that experienced during the 2008-2009 crisis.

Figure 6.4.2.1: Sensitivity Testing Results – Modeled Stress vs. Historical Stress – CMI

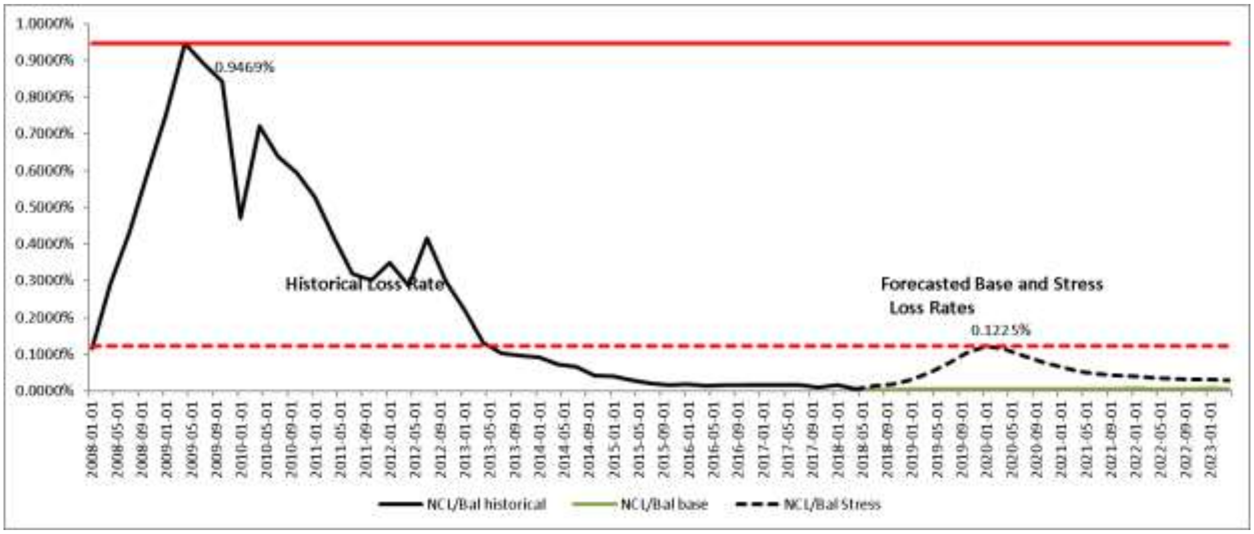


Figure 6.4.2.2: Sensitivity Testing Results – Forecasted Base & Stress vs. Historical Stress –CMI Only

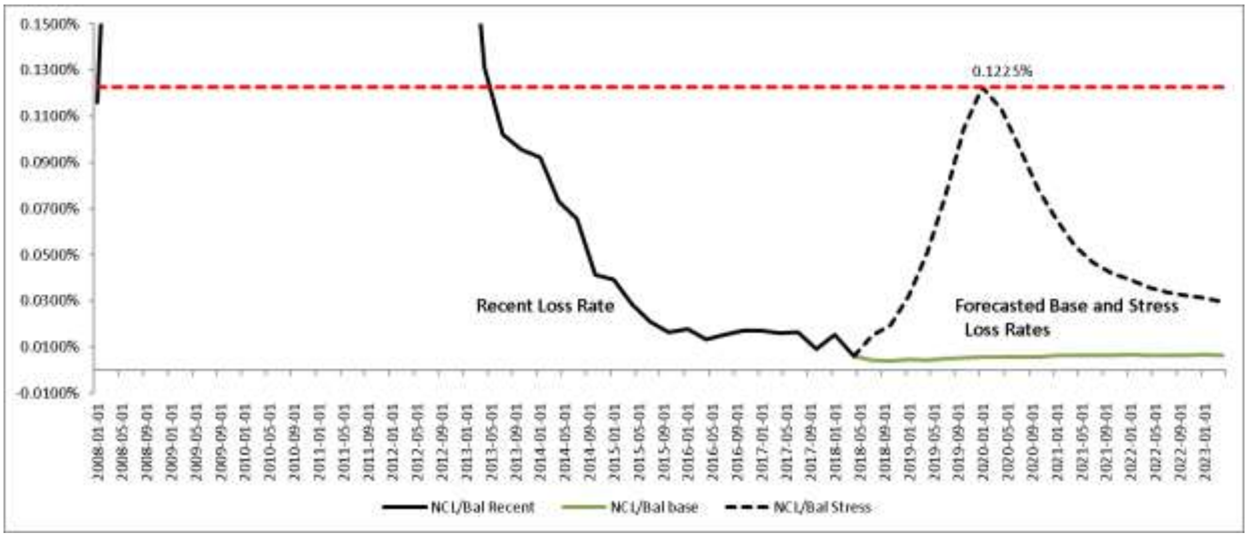
	Max of Forecasted term	MAX Historical	MAX Sensitivity
Base	0.0068%	0.95%	140X
Stress	0.1225%		7.8X

Please note, that a raw comparison of the historical stress and the forecasted stress would not yield economically intuitive interpretation. The comparison has to be made in context of the significant portfolio change post-crisis experience. Today's mortgage portfolio is vastly different to that of a decade ago. As already narrated in Section 1.2 of the MDD, in the after-math of the crisis, Bank's management made sweeping changes to tighten the underwriting standards with a focus on retail channel originations. Further a new entity was created (Bank Holdings) for the strategic preclusion of all non-performing / low quality loans originated during or pre-crisis. Both of these initiatives resulted in a stronger portfolio with improved credit quality. Second, substantial FFIEC write-down during and shortly after the historical stress period, and the strong macroeconomic environment in recent times spurred an increase in housing prices, which coupled with declining unemployment led to significantly lower losses in current times.

Thus, the difference between the modeled Stress scenario and historical stress experience is not so much a reflection of poor model performance or macroeconomic sensitivity but instead the results of portfolio composition refinement, active portfolio risk management and historical FFIEC partial write-down.

In order to better represent the impact of the modeled stress/base scenarios relative to historical stress, CAMU presents Figure 6.4.2.3, focusing on near-term results. The chart provides a visual representation of the model's separation across scenarios in the recent periods. As illustrated, the forecasted stress rates are very closely aligned which attests to the robustness of the model in forecasting future stress episodes. Please note that the forecast starts beginning July 2018. CPB specific figures can be found in attachment - '6.3_6.4_RM.doc.x'.

Figure 6.4.2.3: Sensitivity Testing Results – Modeled Stress vs. Historical Stress, Near Term



Presented below is the 'R2/R1 = R' ratio calculation for the CMI RM portfolio. This ratio has been a new addition in this year's model performance summary results based on MRM's Testing Guidance requirements (release dated Sept 2018). The R ratio provides an assessment of in terms of whether the model predicted losses are in line with the historical losses in the context of the modeled stress vs. historical macroeconomic environments. The threshold for R is set at [50%, 150%]. Any breach of the threshold limits would imply that the modeled stress scenario (incorporating the stressed macro-economic factors) is not sufficiently sensitive enough to capture the historical stress conditions. Ratios that lie within the threshold limits imply that the modeled stress adequately captures the losses similar to the historical stress. The CMI RM portfolio displays of ratio which is N/A which implies there is quite a significant difference between the modeled stress and the historical stress event. As iterated within the initial paragraphs of the MDD, the difference between the modeled stress scenario and historical stress experience is not so much a reflection of poor model performance or macroeconomic sensitivity but instead the results of portfolio composition refinement, active portfolio risk management and historical FFIEC partial write-down. As such, CAMU also conducted an as-if analysis to show how the current portfolio would behave given the historical macroeconomic-conditions from the 2008-2009 stress. Please refer to the as-if analysis presented below.

Table 6.4.2.1a: Sensitivity Testing Results – Modeled Stress vs. Historical Stress, R2/R1 Ratio (CMI Only)

		NCL/ Balance
A	9Q Recent Period Average [**non-stress period**]	0.01%
B	9Q Stress Period Average [***historical stress period***]	0.71%
C	R1 (= B/A)	5091.35%
D	9Q Baseline Forecast Average	0.00%
E	9Q Severely Adverse Forecast Average	0.07%
F	R2 (= E/D)	N/A
G	R = (R2/R1)	**N/A**
H	Result (Pass when R is between 50% to 150%)	**N/A**

The R2/R1 ratio is also calculated for the CPB portfolio too. Please refer to attachment '6.3_6.4.doc' for CPB specific historical to stress and the recent term comparisons. Please note the R2/R1 ratio cannot be estimated for the CPB portfolio[R1 and R2 = 0 for CPB). The results for CPB portfolio are not unsurprising given the lack of sensitivity of the CPB clientele to macroeconomic conditions.

Table 6.4.2.1b: Sensitivity Testing Results – Modeled Stress vs. Historical Stress, R2/R1 Ratio (CPB Only)

		NCL/ Balance
A	9Q Recent Period Average [**non-stress period**]	0.00%
B	9Q Stress Period Average [***historical stress period***]	0.00%
C	R1 (= B/A)	N/A
D	9Q Baseline Forecast Average	0.01%
E	9Q Severely Adverse Forecast Average	0.04%
F	R2 (= E/D)	607.33%
G	R = (R2/R1)	**N/A**
H	Result (Pass when R is between 50% to 150%)	**N/A**

CAMU also conducted "as if" analysis by running the 201806 portfolio *as of* the Jan-2008 historical stress scenario. Table 6.4.2.2 below illustrates the summarized run results for the overall, CMI and CPB portfolios respectively.. Additional details can be found in attachment - '6.4.2 As-if Analysis.xlsx'.

Table 6.4.2.3 – As –if analysis results

RM	Loan Count	Cum IVP units	Cum NCL ($MM)	Quarterly-Max Loss Rate	Quarterly-Avg Loss Rate
RM Overall					
201806+27m C1 (base)	170,274	2,528	32,551,587	0.0136%	0.0057%
201806+27m C8 (Stress)	170,274	4,731	327,278,390	0.1023%	0.0597%
As-if 200801+27m using cum FFIEC at 1806	170,274	5,261	568,273,753	0.1375%	0.0985%
As-if 200801+27m using cum FFIEC at 0801	170,274	5,261	624,812,406	0.1426%	0.1079%
RM CMI					
201806+27m C1 (base)	148,524	2,501	17,234,343	0.006%	0.005%
201806+27m C8 (Stress)	148,524	4,644	237,906,093	0.122%	0.070%
As-if 200801+27m using cum FFIEC at 1806	148,524	5,167	459,908,522	0.178%	0.126%
As-if 200801+27m using cum FFIEC at 0801	148,524	5,167	506,539,684	0.185%	0.139%
RM CPB					
201806+27m C1 (base)	21,750	26	15,317,244	0.030%	0.007%
201806+27m C8 (Stress)	21,750	87	89,372,297	0.069%	0.043%
As-if 200801+27m using cum FFIEC at 1806	21,750	94	108,365,231	0.069%	0.051%
As-if 200801+27m using cum FFIEC at 0801	21,750	94	118,272,722	0.070%	0.055%

Row # 1 represents the 'base' forecast using the 201806 snapshot. Row # 2 represents the 'modeled stress' forecast for the 201806 snapshot. Row # 3 represents the first as-if analysis that utilizes the macro-economic conditions that were prevalent during the 200801 historical stress period while still using the 201806 snapshot information and accompanying portfolio composition and finally, Row # 4 represents the second as-if run on the 201806 portfolio that combines the 200801 historical stress conditions along with each loan's FFIEC write down as of Jan-2008.

Row 1 vs 2 represents the base vs stress differential on today's portfolio mix. The recoveries made from the massive FFIEC write-downs in the past, coupled with rising housing prices, declining unemployment and steadily shrinking MOD volume, has resulted in net negative losses in the base forecast in recent times; which is accurately captured by the model.

Row # 3 represents the effect of overlaying the historical macro-economic stress conditions on today's portfolio mix. As expected, the loss forecast is lower under the 'modeled stress' compared to the 'historical stress' since the macro-economic conditions are way better in recent times (with declining unemployment, rising housing prices, income and overall GDP, positive equity in home with increasing consumer confidence) compared to the historical stress period.

Row # 4 leverages the historical stress experience with historical FFIEC data, which substantiates the significant FFIEC write-downs that underwent during and shortly after the historical stress period as part of a broader companywide initiative to holistically manage the risks and limit loss exposures. As such, the difference between the forecasted 'Stress' loss and 'historical stress' experience is not so much a reflection of lack of macroeconomic sensitivity but instead is an amalgamation of improved portfolio composition, consistently strong home price appreciation, strong macro-economic outlook and significant historical FFIEC write-downs.

RM MEA Dated 10/17/2018 –

Question - Model sponsor attributes the difference between stress forecast with historical maximum loss to changes in portfolio composition. Please provide justification evidence of portfolio composition changes w.r.t the key drivers.

Answer - We illustrate this improved credit quality in Figures 6.4.2.4 and 6.4.2.5 showing the distribution of CLTV and FICO scores, respectively. The distribution of current CLTVs is skewed decidedly in the [00-60] band for the current portfolio relative to that of the 2008 portfolio, reflecting the increased equity in recent times due to consistently improving home prices. Similarly, the distribution of refresh FICO is skewed decidedly higher[750,850] in the current portfolio relative to that of the 2008 portfolio, reflecting higher credit borrowers. Individually and collectively, these comparisons indicate significantly lower credit risk in the recent portfolio compared to the portfolio mix during the stress period. The stronger credit portfolio composition mix is recent times is an amalgamation of strategic sale of lower quality loans, stringent underwriting criteria and improving macro-economic conditions reflected through rising GDP, lower unemployment and increasing housing prices.

Figure 6.4.2.4: Portfolio Composition – Current CLTV

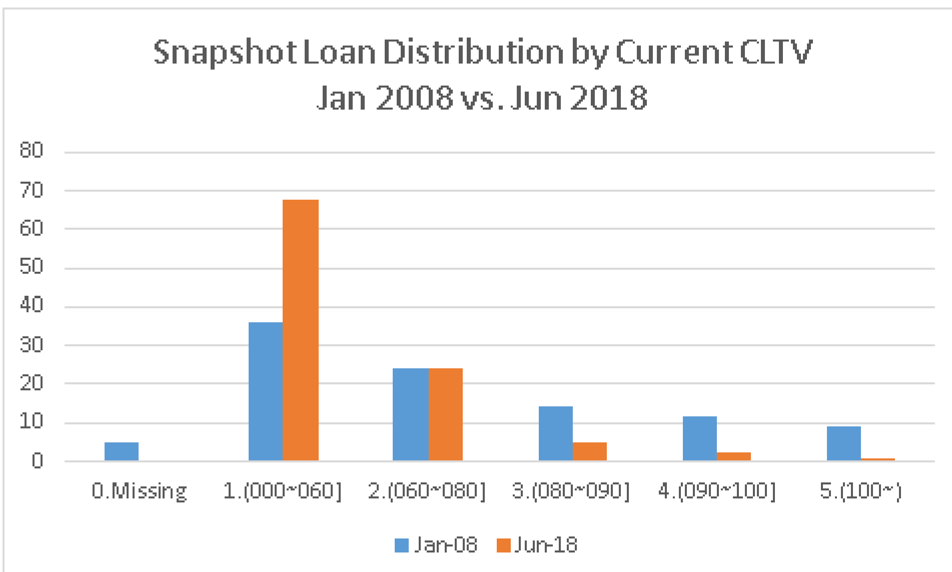
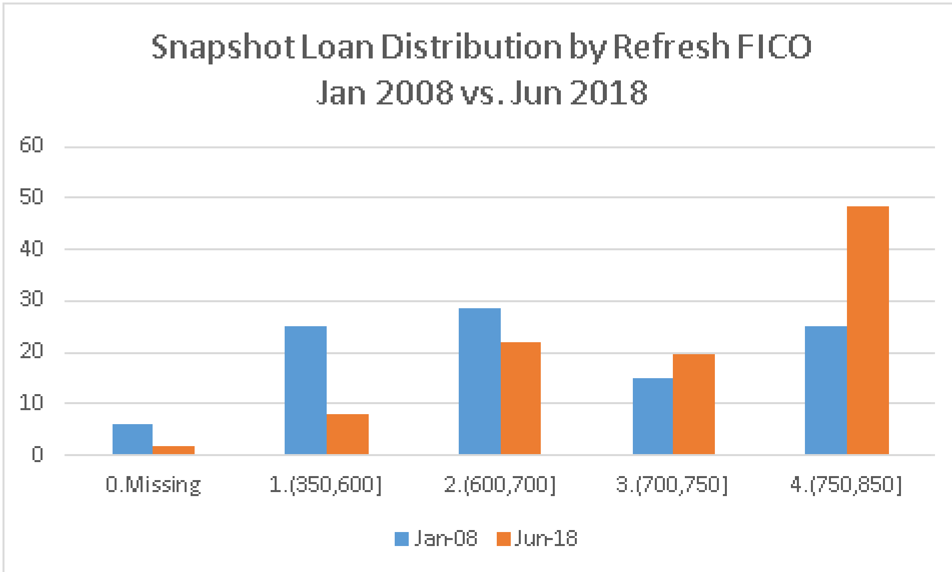


Figure 6.4.2.5: Portfolio Composition – Refresh FICO



Have sensitivity tests been performed at various levels of the model - at the model component level (for e.g. PD, LGD, EAD), as well as final output (for e.g. NCL, Balance)?

> [Describe which components were tested and justify if any of the components were omitted.]
>
> The sensitivity tests have been performed for the NCL metric. The NCL metric captures the model end-to-end performance and hence is considered to provide a holistic comparison of the historical and forecasted stress scenarios. As such, the base/stress scenario comparison with historical stress experience focuses on the NCL rate.

What are the sensitivity test results?

> [Evaluate sensitivity metrics as defined in the Model Testing Guidance for hypothetical stress scenario vs historical stress comparisons. All the results must be justified and their business interpretation provided in the form of the following summary table.]

As discussed above and illustrated in Table 6.4.2.1, loss rate projections in the Stress scenario are far lower than those experienced during the financial crisis. This is attributable to changing portfolio composition, in the after-math of the historical crisis.

Table 6.4.2.4.: Summary of Sensitivity Testing Results - Historical Stress Episode

Model component	Test / shock description	Test results (metric applied and result)	Business interpretation	Mitigation / limitation if performance is not deemed sufficient
GCL rate	Historical vs Modeled Stress	CMI	Historical Stress differs from the Forecasted Stress	Not Applicable. The performance difference can be attributed to significant portfolio composition change and lower losses in recent times.
NCL rate	Historical vs Modeled Stress, Near Term	CMI	Forecasted Stress is adequately sensitive	Not Applicable

6.4.3 Sensitivity of model results to changes in underlying macro variables (univariate and multivariate analysis)

[Describe the design of the sensitivity tests performed perturbing one or multiple underlying variables. The tests should cover comparison of base and stress values of the underlying variables. Both univariate and multivariate tests should be considered. See Model Testing Guidance for details.

Have sensitivity tests been performed at various levels of the model - at the model component level (for e.g. PD, LGD, EAD), as well as final forecasted variable (for e.g. NCL, Balance)?

[Describe which components were tested and justify if any of the components were omitted.]

Yes, sensitivity tests have been performed at the model component level reflecting the PD model's performance through the GCL metric(AMT) as well as the end-to-end model performance through the final forecasted variable – NCL(AMT).

Beyond sensitivity to the overall macroeconomic environment, as demonstrated in Sections 6.4.1 and 6.4.2, the Method A RM Model demonstrates sufficient sensitivity to individual macroeconomic components: GDP, interest rates, income, home prices, and unemployment. An additional macroeconomic attribute that was added this year corresponded to the S&P500 Index. The S&P500 is an American stock market index based on the market capitalizations of 500 large companies and is an excellent leading indicator of the equity market health. Since the equity market performance is closely tied to the overall credit market performance, it was considered prudent to include this attribute as part of the 2019 modeling process.

Have all univariate and multivariate tests been performed?

> [Describe the univariate and multivariate tests performed. If some of the recommended tests are omitted, explain the reasons and list model performance limitations due to missing sensitivity tests.]

Presented in Table 6.4.3.1 is the model's sensitivity results for the univariate and multivariate (Stress) test results for the twenty-seven(27) month forecasting horizon for the CMI portfolio. The 'Parameter Stress' reflects the univariate effect of the stress non-modeled transitions rate (Please refer to Sections 5.1.2 and 3.1 for additional details). Please note all individual add on effect of the all macro-economic parameters are based off the 'Parameter Stress' Setting. In the following table, the row "Stress(ALL)" shows the cumulative difference in the GCL/ NCL amount from the base scenario to the stress scenario throughout the forecast horizon by stressing all macro-level factors on top of the 'Parameter Stress' representing the effect of multivariate sensitivity. In row "IR", the forecast is generated holding the base scenario forecast for all macro-factors except the Interest Rate (IR) factor and Parameter Stress, which are the only two stressed factors. By reviewing these numbers, CAMU concludes that HPI – 1.5X for GCL and 4.33X for NCL has the dominant effect on the model's overall sensitivity given the estimated coefficients of these variables and the forecast scenarios used for testing. This conforms with economic intuition as HPI affects borrower marked-to-market LTV(Loan-to-value) through improved property valuation which in turn reduces all incremental future losses. Mortgages are also quite sensitive to unemployment rate, equity market performance and income, as can be seen in the Table 6.4.3.1 below. Table 6.4.3.2 shows the CPB specific results. For the CPB portfolio, Unemployment Rate is the dominant factor for GCL while HPI is the dominant factor for NCL. This aligns with business intuition as CPB portfolio is not sensitive to macro-economic environment.

Table 6.4.3.1: Sensitivity Testing Analysis of Macroeconomic Factors – CMI

201806 Snapshot 27 months				
Stressed Factor	**Predicted GCL ($MM)**	**Ratio vs. Parameter Stress**	**Predicted NCL ($MM)**	**Ratio vs. Parameter Stress**
Base	592.77		17.23	
Parameter Stress	809.19	1.00	34.76	1.00
Univariate - IR	796.46	0.98	32.81	0.94

Univariate - HPI	1,217.08	1.50	150.44	4.33
Univariate - UR	1,010.42	1.25	61.11	1.76
Univariate - GDP	809.37	1.00	34.79	1.00
Univariate - SP500	809.64	1.00	35.36	1.02
Univariate - Income	850.35	1.05	37.91	1.09
Stress (ALL)	1,572.53	1.94	237.91	6.84
IR + 2%	820.14	1.01	37.31	1.07
IR + 4%	827.28	1.02	39.10	1.12

Table 6.4.3.2: Sensitivity Testing Analysis of Macroeconomic Factors - CPB

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Parameter Stress	Predicted NCL ($MM)	Ratio vs. Parameter Stress
Base	117.21		15.32	
Parameter Stress	234.85	1.00	32.09	1.00
IR	229.58	0.98	31.50	0.98
HPI	263.79	1.12	60.77	1.89
UR	321.04	1.37	40.73	1.27
GDP	234.88	1.00	32.10	1.00
SP500	235.28	1.00	32.24	1.00
Income	251.21	1.07	33.86	1.05
Stress (ALL)	386.61	1.65	89.37	2.78
IR + 2%	237.52	1.01	32.41	1.01
IR + 4%	238.60	1.02	32.51	1.01

Presented below are Tables 6.4.3.3 and 6.4.3.4 which illustrate the incremental GCL/NCL sensitivity by stressing each parameter/macro-economic attribute individually one at a time to reach the cumulative GCL/NCL impact for the CMI and CPB RM portfolios respectively.

Table 6.4.3.3: Incremental Sensitivity Analysis – CMI

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Incremental GCL ($MM)	Predicted NCL ($MM)	Incremental NCL ($MM)
Base	592.77		17.23	
+Parameter Stress	809.19	216.42	34.76	17.53
+IR	796.46	(12.73)	32.81	(1.95)
+HPI	1197.08	400.61	146.69	113.88
+UR	1503.74	306.66	223.30	76.61
+GDP	1503.94	0.21	223.34	0.04
+SP500	1502.58	1.37	223.22	0.12
+Income (ALL Stressed)	1572.53	69.96	237.91	14.68

Table 6.4.3.4: Incremental Sensitivity Analysis – CPB

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Incremental GCL ($MM)	Predicted NCL ($MM)	Incremental NCL ($MM)

Base	117.21		15.32	
+Parameter Stress	234.85	117.65	32.09	16.78
+IR	229.58	5.27	31.50	0.59
+HPI	258.71	29.13	59.64	28.14
+UR	357.41	98.71	82.89	23.25
+GDP	357.44	0.03	82.90	0.01
+SP500	357.77	0.33	82.96	0.06
+Income (ALL Stressed)	386.61	28.84	89.37	6.41

What are effects of the tested macro factors on the model dependent variable and the final forecasted variable (linear, non-linear (concave, convex, spline), additive across macro variables etc.)?

> [Describe the effects of macro factors on the model dependent variable and the final forecasted variable and provide rationale for sensitivity across key range of values of the key macro factors s]

Figures 6.4.3.5 and 6.4.3.6 reflect the walk for GCL and NCL projections from the Base to the Stress using the univariate effects for the CMI portfolio. The 'Parameter Stress' reflects the univariate effect of the stress non-modeled transitions rate. Since the non-modeled rates (Please see Sections 3.1 and 5.1.2 for detailed discussion on non-modeled assumption) are essentially considered model assumption, per MRM's Model Testing Guidance, CAMU has included this assumption as part of its sensitivity-testing grid and within the base-to-stress "walk". Please see Section 6.4.4 for additional details on the non-modeled transition assumption testing. Note that HPI followed by Unemployment rate are the most dominant factors for GCL. The 'Parameter Stress' has a relatively smaller effect compared to HPI and Unemployment for GCL. Similar direction and magnitude can be observed for the CMI NCL Walk too.

Figure 6.4.3.5: Macroeconomic Sensitivity – Walk from Base to Stress (GCL) - CMI

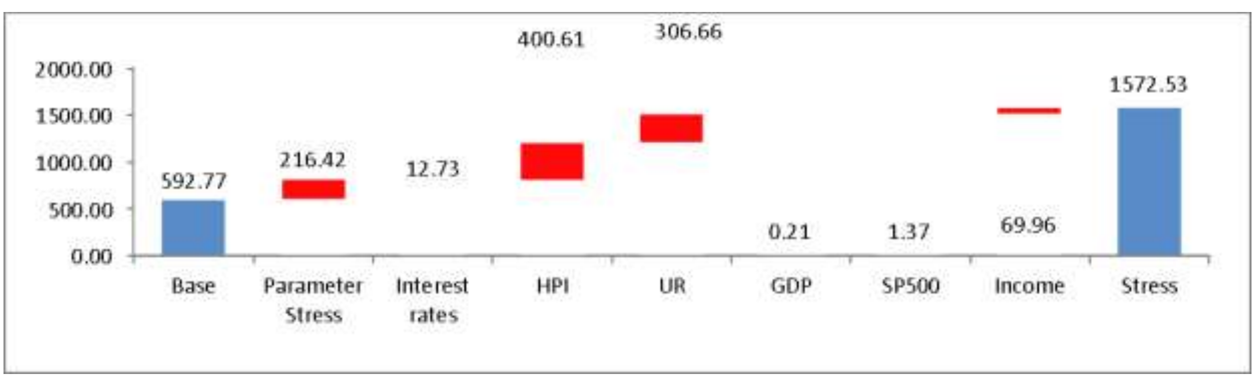
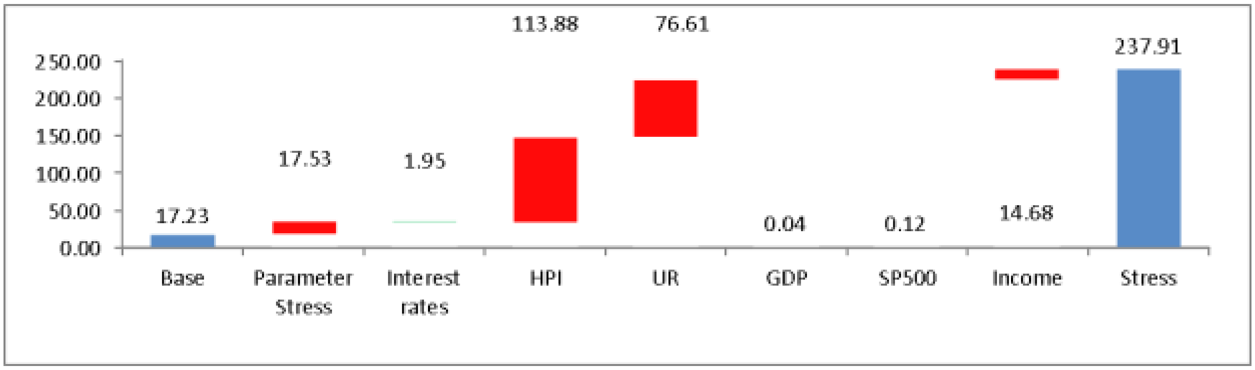


Figure 6.4.3.6: Macroeconomic Sensitivity – Walk from Base to Stress (NCL) - CMI



Presented below are Figures 6.4.3.3 and 6.4.3.4 which illustrates the GCL/NCL Walk test results for the CPB portfolio for the twenty-seven month horizon. For CPB - the 'Parameter Stress' setting has the dominant effect followed by Unemployment rate for the GCL. For NCL, HPI and Unemployment Rate have the dominant effects compared to Parameter Stress Setting.

Figure 6.4.3.3: Macroeconomic Sensitivity – Walk from Base to Stress (GCL) – CPB

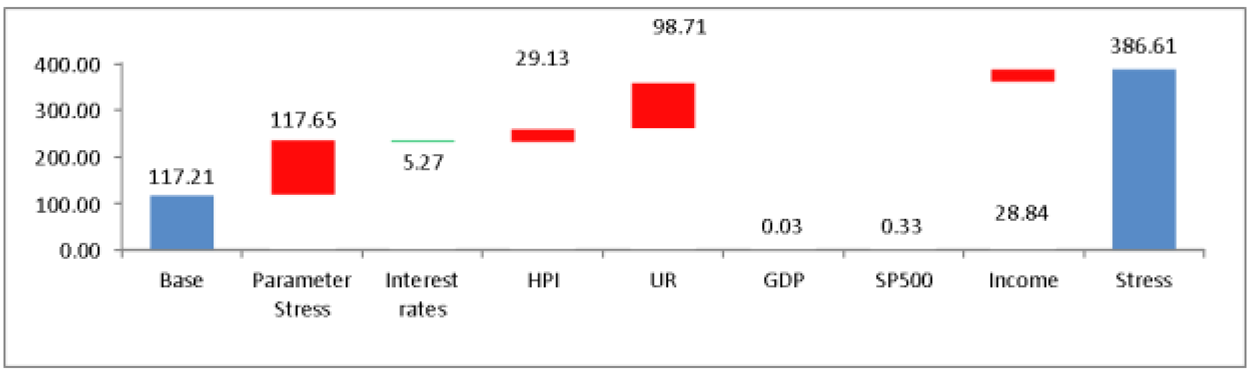
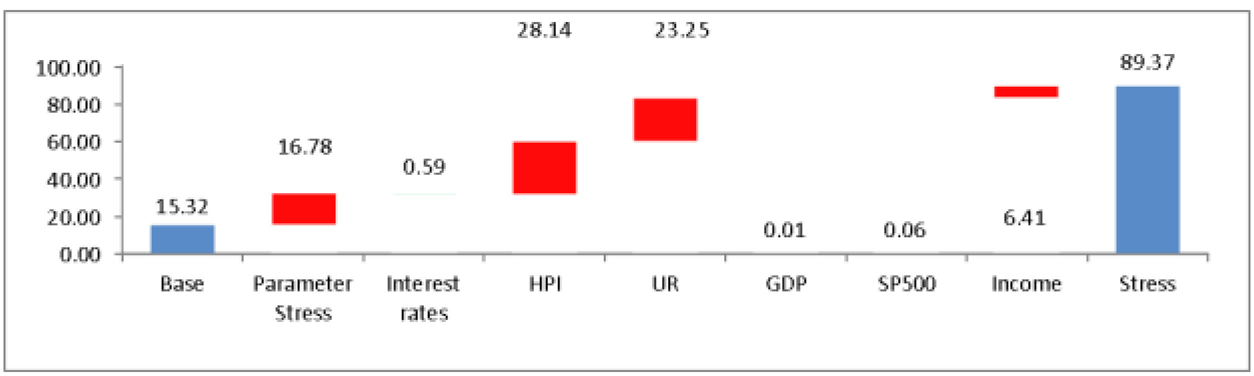


Figure 6.4.3.4: Macroeconomic Sensitivity – Walk from Base to Stress (NCL) – CPB



What variables and values were tested?

[For all sensitivity tests performed, describe the data used, shocks in variables considered and model components, if applicable, that were tested (testing impact on NCL and GCL characteristics is the prescribed minimum).]

[If additional sensitivity tests were performed, one should describe and explain the model component tested, the data used, the assumptions that were applied, the business rationale for the tests, and the formulas used to evaluate the results.]

As discussed above, CAMU tested the effects of interest rates, income, GDP, S&P500 index, home prices, and unemployment on model projections of GCL and NCL amount for both CMI and CPB portfolios separately.

What are the sensitivity test results?

[Evaluate sensitivity metrics as defined in the Model Testing Guidance for non-stress and stress comparisons. All the results must be justified and their business interpretation provided in the form of the following summary table.]

[Describe test results and provide the rationale for whether results are reasonable and expected. Explain why or why not.]

[For multivariate scenarios, provide waterfall analysis of impacts as illustrated in the Model Testing Guidance. Apart from the Excel table requested above, provide any additional files needed to document sensitivity testing.]

Presented below is the summarized Table of model results. Macroeconomic sensitivity tests confirm intuition.

Table 6.4.3.3: Summary of Sensitivity Testing Results - Macro Variables Sensitivity

Model component	Test / shock description	Test results (metric applied and result)	Business interpretation	Mitigation / limitation if performance is not deemed sufficient
GCL/NCL	27 month forecast	Univariate/Multivariate Analysis for CMI Portfolio	Model is sufficiently sensitive to macro-economic attributes	Not Applicable
GCL/NCL	27 month forecast	Univariate/Multivariate Analysis for CPB Portfolio	Model is sufficiently sensitive to macro-economic attributes	Not Applicable

6.4.4 Sensitivity of model results to changes in model parameters

[Examine the sensitivity of the model to changes in the values of model parameters which are set based on business assumptions and deemed to depend on macroeconomic conditions.

Describe model parameters to be tested and the shocks applied. Provide business justification of the shocks considered. Perform univariate tests. If applicable, consider multivariate shocks as well. See Model Testing Guidance for details.

As mentioned within Sections 3.1 and 5.1.4, the Method A RM model segregates the all of the PD transitions into modeled vs non-modeled transitions. While modeled transitions are modeled using regression equations, non-modeled transitions utilize an empirical time-varying rate which takes into account the recent vs stress differential, government vs conventional and trial vs not-in-trial.

DV Haircut rates and a CPB severity multiplier are other parameters used in the model. These parameters are both described within corresponding 5.1.2 sub-sections.

Two types of sensitivity tests performed to evaluate model sensitivity to parameter changes and shocks are described below.

6.4.4.1 Sensitivity to Parameter Stress Setting

The first test is a parameter stress setting test to compare model results under a base macroeconomic scenario with the parameter stress setting OFF and then ON.

The ‘Parameter Stress’ row in Table 6.4.4.1.1 reflects sensitivity of model results to Parameter Stress Setting vs. Non-stress setting. Sensitivity testing has been done on the PD (GCL) and the final output (NCL) metrics. The results indicate that the model is sufficiently sensitive to stress scenario parameter settings used in the model. Similar testing has been executed on the CPB portfolio too. The test results for CPB also demonstrate similar directional change.

Table 6.4.4.1.1: Effect of switching model parameters to stress setting – CMI

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	592.77	1.00	17.23	1.00
Parameter Stress	809.19	1.37	34.76	2.02
Stress	1,572.53	2.65	237.91	13.80

Table 6.4.4.1.2: Effect of switching model parameters to stress setting – CPB

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	117.21	1.00	15.32	1.00
Parameter Stress	234.85	2.00	32.09	2.10
Stress	386.61	3.30	89.37	5.83

6.4.4.2 Sensitivity to Non-Modeled Parameter Shock

To examine PD and NCL result changes from shocks to non-modeled transition rates, a 10% shock rate was applied on these rates. Specifically, jump-to-worse transition base rate was inflated by 10% while jump-to-cure transition base rate was deflated by 10%. Based on the results described above, the model has marginal responsiveness to shocks on non-modeled transition rates. As shown in the table, shocking non-modeled transitions by 10% slightly increased both GCL and NCL.

Table 6.4.4.2.1 Effect of shocking non-modeled transition rate by 10% - CMI

201806 Snapshot 27 months				
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base	Predicted NCL ($MM)	Ratio vs. Base
Base	592.8	1	17.2	1
shocked non-model	604.5	1.02	18.0	1.05

Table 6.4.4.2.2 Effect of shocking non-modeled transition rate by 10% - CPB

201806 Snapshot 27 months				
Stressed Factor	**Predicted GCL ($MM)**	**Ratio vs. Base**	**Predicted NCL ($MM)**	**Ratio vs. Base**
Base	117.2	1	15.3	1
shocked non-model	121.0	1.03	15.7	1.03

Have sensitivity tests been performed at various levels of the model -- at the model component level (for e.g. PD, LGD, EAD), as well as final output (for e.g. NCL/ Balance)?

[Describe which components were tested and justify if any of the components were omitted.]

Sensitivity testing conducted on the PD non-modeled transitions has been discussed in section 6.4.4.2 above.

What model parameters and values were tested?

[For all sensitivity tests performed, describe the data used, shocks in model parameters considered and model components, if applicable, that were tested (testing impact on NCL and GCL characteristics is the prescribed minimum). See an illustration below. Provide a similar table for the tested model.]

Sensitivity testing conducted on the PD non-modeled transitions has been discussed in section 6.4.4.2 above.

Table 9: Sensitivity Testing -- Variables and Scenarios Considered in Sensitivity to Model Parameters Testing

Scenario	**Base**		**Stress**	
Parameter / Value	**10%**	**-10%**	**10%**	**-10%**
Non-Modeled Transitions	See discussion in section 6.4.4.2 above.			

[If additional sensitivity tests were performed describe and explain the model component tested, the data used, the assumptions that were applied, the business rationale for the tests, and the formulas used to evaluate the results.]

See discussion in section 6.4.4.2 above.

What are the sensitivity test results?

[Evaluate sensitivity metrics as defined in the Model Testing Guidance for non-stress and stress comparisons. All the results must be justified and their business interpretation provided in the form of the following summary table.]

Table 6: Summary of Sensitivity Testing Results -- Sensitivity to Model Parameters

Model component	Test / shock description	Test results (metric applied and result)	Business interpretation	Mitigation / limitation if performance is not deemed sufficient
See discussion in section 6.4.4.2 above.				

[Describe test results and provide the rationale for whether results are reasonable and expected. Explain why or why not.]

[For multivariate scenarios, provide waterfall analysis of impacts. Attach supporting analytical files with all relevant results.]

See discussions within sub-sections 6.4.4.1 and 6.4.4.2 above

6.4.5 Sensitivity of model results to changes in the key assumptions

[The modeler should include results of impact on model results based on changing the key business assumptions or inputs.]

Have sensitivity tests on assumptions been performed at various levels of the model - at the model component level (for e.g. PD, LGD, EAD), as well as at the level of model sub-components (such as vintage quality assumptions for PD in case of Look Ahead models, etc.)?

[Describe which assumptions were tested and justify if any of the components were omitted.]

This is not applicable for the Method A RM PD or LGD model.

Have all the relevant business assumptions or inputs been tested? What key assumptions and values were tested?

[Describe key business assumptions or related model inputs to be tested and the shocks applied. Provide business justification of the shocks considered. Perform univariate tests. If applicable, consider multivariate shocks as well.]

[For all sensitivity tests performed, describe the data used, shocks in key assumptions considered and model components, if applicable, that were tested (testing impact on NCL and GCL characteristics is the prescribed minimum). See an illustration below. Provide a similar table for the tested model.]

Example Table 7: Summary of Original and Shocked Assumption Descriptions

Scenario	Base		Stress	
Assumptions / Shocks	**Original**	**Shock**	**Original**	**Shock**
Assumption 1				
Assumption 2				
Assumption 3				
Assumption 4				
Assumption 5				

[If additional sensitivity tests were performed describe and explain the model component tested, the data used, the assumptions that were applied, the business rationale for the tests, and the formulas used to evaluate the results.]

Business assumptions, for instance future origination quality, are related to post model application. These types of business assumptions are further analyzed for potential inclusion in the overlay analysis and documentation process.

What are the sensitivity test results?

[Please provide results of assumptions sensitivity testing. All the results must be justified and their business interpretation provided in the form of the following summary table.]

Example Table 8: Summary of Sensitivity Testing Results -- Assumptions Sensitivity

Model component	Test / shock description	Test results	Business interpretation

[Describe test results and provide the rationale for whether results are reasonable and expected. Explain why or why not.]

Not Applicable. See discussion above

6.4.6 Sensitivity of model results – SFO Portfolio

Since the Method A RM model is borrowed for SFO usages (Please refer to Section 1.1 for the Model Usage Grid), CAMU had conducted an additional set of sensitivity testing, specific to the SFO RM portfolio only. Presented below are the sensitivity results. Please note that all results below are for GCL only. NCL is not a relevant metric for the SFO portfolio.

Table 6.4.6.1: Summary of Sensitivity Testing Results – Base vs. Stress

201806 Snapshot 27 months		
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base
Base	565.45	1.00
Stress	1,530.03	2.71

Table 6.4.6.2: Sensitivity Testing Analysis of Macroeconomic Factors

201806 Snapshot 27 months		
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Parameter Stress
Base	565.45	
Parameter Stress	811.30	1.00
IR	797.40	0.98
HPI	1,161.62	1.43
UR	1,009.89	1.24
GDP	811.51	1.00
SP500	811.99	1.00
Income	860.04	1.06
Stress (ALL)	1,530.03	1.89
IR + 2%	822.63	1.01
IR + 4%	825.38	1.02

Table 6.4.6.3: Incremental Sensitivity Analysis – Walk from Base to Stress (GCL)

201806 Snapshot 27 months		
Stressed Factor	Predicted GCL ($MM)	Incremental GCL ($MM)
Base	565.45	
+Parameter Stress	811.30	245.85
+IR	797.40	(13.90)
+HPI	1144.23	346.83
+UR	1444.03	299.80
+GDP	1444.27	0.24
+SP500	1442.90	1.37
+Income (ALL Stressed)	1530.03	87.13

Table 6.4.6.4 : Parameter Sensitivity – Non-Modeled Transition Shock

201806 Snapshot 27 months		
Stressed Factor	Predicted GCL ($MM)	Ratio vs. Base
Base	565.45	1.00
Parameter Stress	811.30	1.43
Stress	1,530.03	2.71

Figure 6.4.6.1: Macroeconomic Sensitivity – Walk from Base to Stress (GCL)

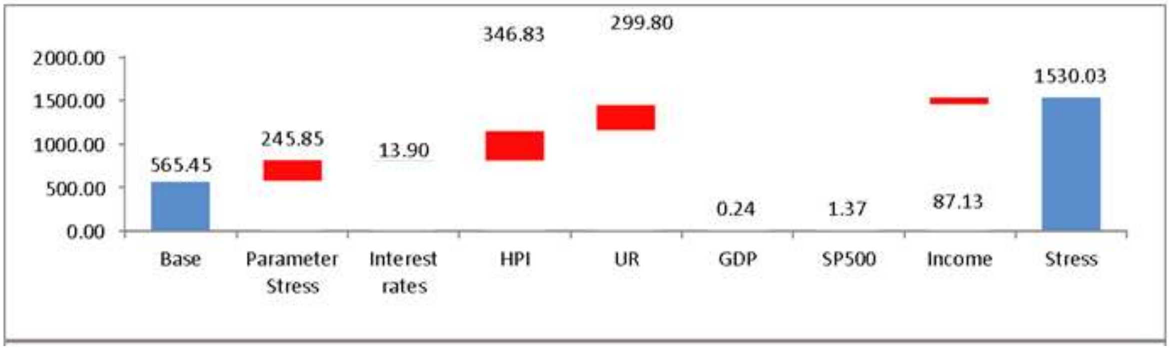


Table 6.4.6.4 : Summary of Sensitivity Results for the SFO RM Portfolio

Model component	Test / shock description	Test results (metric applied and result)	Business interpretation	Mitigation / limitation if performance is not deemed sufficient
GCL	Base vs Stress	27 month	Results are sufficiently sensitive across base and stress	Not Applicable
GCL	Modeled Stress vs Historical Stress	27 month	Modeled Stress is not as sensitive compared to Historical Stress	Portfolio Composition Change
GCL	Univariate / Multivariate Analysis	27 month	Results are sufficiently sensitive for all individual/multivariate macro-factors	Not Applicable
GCL	Parameter Sensitivity	27 month	Results are sufficiently sensitive to shocks in the parameter assumption	Not Applicable

6.5 Benchmark Model Results and Other Triangulation Analyses

[Model Sponsor must provide complete details on the benchmarking analysis including:

- The details of the benchmark model.
- Rationale for choosing a particular model as benchmark.
- Objective and high level modeling framework of the benchmark model
- Results comparison and commentary on the results
- Codes and datasets to replicate the benchmark results

Benchmarking results must be performed in terms of Backtesting and Scenario Forecasting in the template as required by the Model Testing Guidance. Additionally, the Model Sponsor is required to provide the complete benchmark details per the model usage grid. In case alternate models are not available for benchmarking purpose, Model Sponsor must justify the rationale for the same and consider alternate benchmarking options as mentioned in the Model Testing guidance. Please refer to benchmarking section of the Model Testing Guidance for detailed requirement on benchmarking and triangulation analysis.]

As per CAMU's Annual CCAR Plan (please see attachment – '6.5 Modeling Plan discussion') that was shared, reviewed and approved by the relevant model stakeholders, the Method A model suite would be benchmarked to the brand new Method C suite. The Method C suite would leverage a completely different modeling framework to mitigate the prior model's limitation around similarity of the modeling framework. Please refer to Limitations Summary in Chapter 2 for pertinent details . Given the newness of the framework surrounding Method C, it has been granted a delayed submission timeline (January 2019). As such, the Method A vs C benchmarking analyses would be conducted post the completion of the Method C validation at the end of quarter one 2019.

As such, all benchmarking analyses would be conducted post the completion of the Method C validation at the end of quarter one 2019.

Additionally, the RM model suite was also benchmarked to prior year's version to evaluate and assess the model's performance given the re-development attempted this cycle. The prior year's version of the model is an appropriate and proper benchmark: as it employs similar model structure. For a list of 2019 re-development items, please refer to Sections 1.1 and 5.2 of the MDD.

CAMU conducted the benchmark analysis between the 2018 Method A 1st Lien Change Addendum[Model ID – 110324/110389] results and 2019 Method A RM model[Model Id – 167125] suite to assess the change in model performance, given the targeted redevelopment efforts for the 2019 model development process.

Presented below are the summarized tables showing the 'CERRPCT' metric comparison between the newly developed model and the prior model suite for GCL Balance (PD model performance) and NCL (end-to-end model performance) over the nine quarter (9) horizon for both stress(200801) and non-stress(201603) snapshots for the CMI and CPB RM portfolios respectively. Also provided are the nine-quarter forecasts for the 201806 snapshot. The 'Actual Model' refers to the 2019 RM Model [Model ID – 167125] results while the 'Benchmark Model' refers to the 2018 Model results. Additional testing has been performed on four(4), twenty(20) quarters and over all intermediate delinquency buckets, IVP, VP(both balances and units) for both the recent and stress periods to align with the Model Usage Grid requirements, as illustrated in Section 1.1. Please refer to attachments – '6.5 RM_Benchmark.xlsx' and '6.5 RM Benchmarking 27 month results.xlsx for additional details.

Table 6.5.1 : Backtest Comparison – Recent(non-stress) Period- CMI

RM CMI	GCL		NCL		GCL error %		NCL error %	
	2019	2018	2019	2018	2019	2018	2019	2018
Actual	1,024	1,021	50	50				
Forecast	1,067	1,166	53	77	4.2%	14.2%	6.8%	54.6%

Table 6.5.2 : Backtest Comparison – Recent(non-stress) Period- CPB

RM CPB	GCL		NCL		GCL error %		NCL error %	
	2019	2018	2019	2018	2019	2018	2019	2018
Actual	118	118	5	5				
Forecast	107	248	7	17	-9.6%	109.8%	23.1%	209.7%

Table 6.5.3 : Backtest Comparison – Stress Period- CMI

RM CMI	GCL		NCL		GCL error %		NCL error %	
	2019	2018	2019	2018	2019	2018	2019	2018
Actual	18,906	18,906	5,943	5,943				
Forecast	17,903	17,973	6,146	5,994	-5.3%	-4.9%	3.4%	0.9%

Table 6.5.4 : Backtest Comparison – Stress Period- CPB

RM CPB	GCL		NCL		GCL error %		NCL error %	
	2019	2018	2019	2018	2019	2018	2019	2018
Actual	226	226	6	6				
Forecast	360	405	63	82	59.4%	79.4%	891.2%	1178.4%

Table 6.5.5 - Sensitivity Comparison – 9 Quarters – Stress/Base Differential – CMI

RM CMI	GCL		NCL		GCL stress/base		NCL stress/base	
	2019	2018	2019	2018	2019	2018	2019	2018
Base	593	810	17	32				
Stress	1,573	1,370	238	281	2.7	1.7	13.8	8.8

Table 6.5.6 - Sensitivity Comparison – 9 Quarters – Stress/Base Differential – CPB

RM CPB	GCL		NCL		GCL stress/base		NCL stress/base	
	2019	2018	2019	2018	2019	2018	2019	2018
Base	117	364	15	17				
Stress	387	672	89	157	3.3	1.8	5.8	9.3

The following pointers summarize the current model's performance over its prior version.

i. Recent period back testing has significantly improved to mitigate the existing model's overestimation. The NCL error over prediction was reduced by almost 88% [54.6% to 6.5%] for CMI and 89% for CPB[209.7% to 23.1%] . Similarly, the GCL error [actual vs predicted] dropped by ~73% for CMI and over 100% for CPB.

ii. The Stress period back test results demonstrated conservatism on NCL metric, which is not surprising given the targeted model re-developments attempted this year.

iii. The forecast sensitivity ratio has increased significantly for the CMI portfolio compared to the existing model, again based on the improvements made to the DV Logic, inclusion of recent period in the development data, leveraging a standardized approach to segregate and estimate the modeled vs non-modeled transitions and enhanced segmentation analysis.

Separate sensitivity testing(201806 snapshot used) has been conducted on the DV Logic used in the 2019 CCAR model. The 2019 DV Logic has been benchmarked to the 2018 CCAR results on the 2017 DVM model[model ID – 108730]. These results have been illustrated in Section 3.2.2.3 of the MDD. Please refer to this section for additional details.

6.6 Model Interdependencies or Interconnectivity

[It is incumbent upon developers to identify cases where there are interdependencies or inter connectivity with respect to other models and perform sufficient testing to quantify, measure, interpret, and document its effects on model performance including nested model errors and model reliability. Some examples of situations when a nested model error may arise include: error due to relying on outputs of other (feeder) models; error due to forecasting of macroeconomic inputs.]

[Identify cases where the model is using output from other models, and provide a brief description of those models along with the Model IDs. Also highlight the stages where these models are used within this model.]

[If this model or its output is used within other models, then provide a description of all such models and how this model is applied within those models. Also provide the Model IDs for all such models. Please describe the testing conducted to calculate nested error. Please also describe the outcome of the testing.

The Method A RM model is subject to nested model error in its reliance on macroeconomic forecasts.

As demonstrated in Section 6.4, this model is sensitive to macroeconomic conditions. As such, forecasts of model outputs (PD and LGD) as well as outcome variables (e.g., GCL and NCL) require forecasts of macroeconomic conditions. However, the future is uncertain; the forecasts developed by GCRM and utilized by CAMU, as well as downstream users, will almost undoubtedly be wrong. Model error in GCRM's upstream macroeconomic models will propagate through the Method A RM model which will then cascade to all downstream model usages.

One common example is a model, which disaggregates a national level home price scenario into state/county specific home price paths for use in the model. In this case nested model error is defined as the incremental error in forecasts due to using predicted state/county level home price inputs rather than actual historical state/county level home price inputs. Therefore, CAMU implements the nested error analysis by just replacing the county level HPI with the US level HPI. In order to quantify this potential error, CAMU compared the twenty-seven (27) month backtesting results for the recent(March 2016 snapshot) and stress period(January 2008 snapshot) with the existing county level HPI to a revised HPI model, one that uses the aggregated national level forecast for the RM PD model. A similar twenty-seven (27) month backtest was conducted for the LGD model based on the recent(March 2016 snapshot) and stress period(January 2008 snapshot. Table 6.6.1 below contrasts the predicted GCL (AMT) and NCL (AMT) using US level HPI and the county HPI projections for the PD model while Table 6.6.2 below contrasts the predicted NCL(AMT) between the two levels of HPI for the LGD model.

The results from the comparative analyses are summarized as below-

Table 6.6.1: Nested Error Analysis – PD Model

Snapshot Dt		Actual	Based on County HPI		Based on US HPI	
			Prediction	Err%	Prediction	Err%
Jan2008 27M	**GCL ($MM)**	$19,131	$18,262	-5%	$18,571	-3%
	NCL($MM)	$5,949	$6,209	4%	$6,075	2%
	NCL/GCL Ratio	0.31	0.34	9%	0.33	5%
Mar2016 27M	**GCL ($MM)**	$1,142	$1,174	3%	$1,129	-1%
	NCL($MM)	$55	$60	8%	$52	-6%
	NCL/GCL Ratio	0.05	0.05	5%	0.05	-5%

Table 6.6.29: Nested Error Analysis – LGD Model

Snapshot	Actual Loss ($MM)	County Level HPI		US Level HPI	
		Predicted Loss ($MM)	Loss CERRPCT	Predicted Loss ($MM)	Loss CERRPCT
Jan-08	10,061	10,284	2.2%	10,075	0.1%
Mar-16	381.9	395	3.4%	388.5	1.8%

i. The nested error analysis shows that using the US level HPI does not significantly impact the overall NCL forecast at the aggregated level for both PD and LGD models.
ii. While the relative error (CERRPCT) difference is noticeable, the absolute levels are quite small and essentially immaterial. The relative error is thus a small denominator problem.
iii. CAMU will continue to utilize the fully granular HPI model, as it introduces no significant biases and better reflects individual loan/property heterogeneity.

Additional analysis can be found in the attachment – '6.6 Residential Mortgage Model PD Nested Error Analysis for March 2016 and January 2008.xlsx' and '6.6 Severity Model Nested Error Analysis for March 2016 and January 2008.xlsx'.

6.7 Other Performance Tests (if Applicable)

[Please describe any other performance tests conducted, including the types of tests, outcomes, and results]

6.7.1 Performance Monitoring

The MDD represents performance testing at a point in time. The utility of a model on a go-forward basis requires an on-going performance-monitoring plan. CAMU satisfies this through formalized quarterly and annual performance reviews, including engagement and discussion of key performance metrics with the model stakeholders, which comprise: the Model owners/users, the CMI Model Governance and Control Team, the Consumer Analytics and Modeling Unit, independent oversight (Model Risk Management/Independent Risk/Senior Modeling Specialist with appropriate entitlements), the CMI Model Risk Governance Committee, and the Model Sponsor.

Performance is tracked and reported quarterly via MIS reporting to all lines of defense and independently validated annually through Model Risk Management review, of which Internal Audit selects a sample of model MDDs for review. Both the quarterly and annual reviews monitor the Model's predictive performance through a comparison of predicted outcomes to actual values for each period.

All such monitoring and tracking, including but not limited to performance measures, threshold definitions, escalation process, and threshold breach protocols, is performed in accordance to MRM Policy, as well as GCCFRP Chapter 17: Global Consumer Risk Management Supplement, specifically Section 8.2.2: Performance Monitoring, Section 8.3.3: Ongoing Implementation Testing, and Exhibit 17H.

6.7.2 Risk Rank Ordering - KS Statistic

As iterated in Section 1.1 of the MDD, the 2019 Method A RM model suite is proposed for many non-CCAR usages, especially with regards to the holistic management of the risk appetite framework(RAF) for the entire NA(mortgage portfolio as set forth in Bank's Global Consumer Credit Fraud Risk Policy (GCCFRP)). All pertinent details around Bank's Risk Appetite Framework (RAF) are provided in the Global Consumer Credit Fraud Risk Policy (GCCFRP) annexure made available in the Appendix section. To conform applicability of the model for all RAF related measures and outputs, the 27 months cumulative

IVP rate has been chosen as a measure of the model's risk rank ordering power. CAMU calculated the KS statistic for 27-month cumulative IVP rate based on the March 2016 snapshot. The KS statistic for IVP is 85.6% as illustrated below, demonstrating strong rank ordering power from the model. Additional details can be found in attachment – '6.7.2 KS Statistics_Risk_Rank Ordering.xlsx.'

Table 6.7.2 K-S Risk Ranking – IVP Rate

LEVEL	TOT_NUM	BAD_NUM	GOOD_NUM	MIN_SCORE	MAX_SCORE	MEAN_SCORE	CUM_BAD_PCT	CUM_GOOD_PCT	CUM_TOT_PCT	KS
1	19,742	1	19,741	0	0.0003	0.0003	0.02%	10.28%	10.00%	0.10
2	19,742	1	19,741	0.0003	0.0004	0.0004	0.04%	20.56%	20.00%	0.21
3	19,742	4	19,738	0.0004	0.0005	0.0005	0.11%	30.83%	30.00%	0.31
4	19,741	3	19,738	0.0005	0.0007	0.0006	0.17%	41.11%	40.00%	0.41
5	19,744	20	19,724	0.0007	0.001	0.0009	0.54%	51.38%	50.00%	0.51
6	19,742	24	19,718	0.001	0.0017	0.0013	0.99%	61.64%	60.00%	0.61
7	19,741	48	19,693	0.0017	0.003	0.0022	1.89%	71.89%	70.00%	0.70
8	19,743	84	19,659	0.003	0.0061	0.0043	3.46%	82.13%	80.00%	0.79
9	19,741	173	19,568	0.0061	0.019	0.0105	6.70%	92.32%	90.00%	0.86
10	19,744	4,988	14,756	0.019	0.9363	0.2217	100.00%	100.00%	100.00%	-

7. Model Implementation

7.1 Implementation Overview

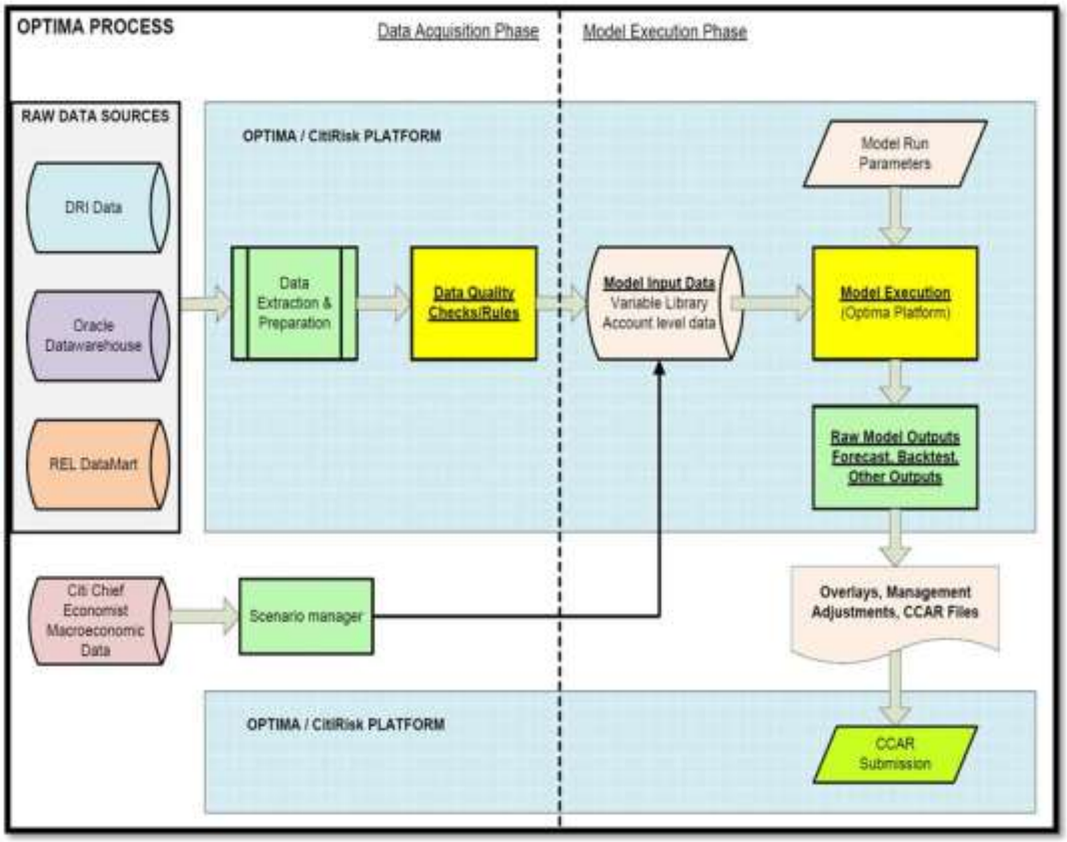
[Describe the environment in which the model will be implemented, and the model implementation execution process.]

As part of model implementation, MRM Policy and GCCRFP Chapter 17 require the business to perform Production Code Review, Pre-implementation Testing (UAT), and Post-implementation Testing (PIT).

As part of model implementation, MRM Policy and GCCRFP Chapter 17 require the business to perform Production Code Review, Pre-implementation Testing (UAT), and Post-implementation Testing (PIT).

The modeling execution process starts by obtaining all required input data. The proposed CCAR model demands two sets of input data, loan characteristic data and macroeconomic data. The loan characteristic data is loaded from the OPTIMA/Model Execution Platform (MEP), which has been sourced from REL DataMart, Oracle Data Warehouse as well as an internal automated default system known as DRI, (provided by a vendor called Default Resolution Inc.). The Macroeconomic data is based on historical Macroeconomic data and future economic expectations/scenarios from Bank's Global Country Risk Management (GCRM) team, CoreLogic, OCC, and Federal Reserve Bank. For each prediction, when the observation month is settled, the loan characteristic data will be established and unchanged, but future economic assumptions can be changed. After the loan characteristic data and the macroeconomic data are processed, they are merged and fed to the variable creation code to create regression ready forecasting input variables. The forecasting code will finally utilize the forecast inputs to generate the Model’s forecast output.

Figure 7.1: MEP OPTIMA PROCESS



The environment in which the model will be implemented, executed, and backed up occurs on Bank's centralized SAS Grid. The development and production code is also uploaded to the MRM Shared Directory as specified by MRM's Code and Data Guidance for purposes of independent review and validation (i.e. model replication).

Specifically, as part of the model implementation phase, the Model Sponsor is responsible for ensuring the following:

- Ensuring that all models reside in a controlled environment where no unauthorized or unintended changes can be made
- Implement appropriate system controls and testing to ensure that appropriate data is used in the proposed model. Segments should establish standards and procedures for data and reporting integrity, and should comply with the Bank Data Management Policy and Standards.
- Models which are released in a production environment are subject to applicable standards for software development and testing (e.g., Bank Solution Delivery Life Cycle Standard (CSDLC)) by the function responsible for the production release (e.g., Front Office IT).
- Post validation, the Model Sponsor must submit evidence to support that a model is implemented in a controlled environment and that the implemented model produces intended results as part of the model development documentation.

It is important to note that models that are implemented as EUCs are subject to the Bank Data Management Policy and Standards, as well as End User Computing (EUC) Standards. Once the model is validated and approved, the Model Sponsor will ensure it has been registered in the Bank EUC Control Portal Inventory.

The Implementation testing plan and detailed steps are further discussed below in MDD section 7.2.

7.2 Implementation Testing

7.2.1 Implementation Testing Plan

[Describe the plan of how implementation will be tested, metrics that will be measured, and the expected outcome of a correct implementation. The implementation testing plan should also include tests the Sponsor plans to perform, the user testing approach, system testing approach etc]

Prior to submission of the implementation addendum for this model, implementation testing is conducted by the SIMPLE (In Business Strategic Initiatives, Model, & Project Lifecycle Execution) Team working along with the Model Code Production team within the Consumer Analytics Modeling Unit (CAMU).

For production code review, the accuracy of the production code calculations was verified through the following steps:

- Checking that data comes from the correct official data source
- Checking that the names and transformations of the variables are correct
- Checking that Model coefficients in the code match to the Model specifications provided by CAMU
- Reviewing the logic in different parts of the code
- Preparing the report and saving the relevant files

In pre-implementation testing, the accuracy of implementation system inputs, calculations, and outputs must be tested prior to implementation. This model is exempt from pre-implementation testing because the model will be implemented using platforms, software programs, and data sources that are the same as those verified during model development code review.

For Post-Implementation Testing (PIT), the accuracy of implementation system inputs, calculations, and outputs are verified through the following steps:

- SIMPLE defines the snap date year and month for the input data.
- SIMPLE will run the code to generate the data, after that SIMPLE will validate result with the data that has been generated by the Model Code Production team.
- SIMPLE will compare two sets of inputs, calculations, and results - Confirm agreement between Model Code Production team & SIMPLE.
- Prepare the report and save the relevant files.

Please refer to the “7.2_Implementation Addendum” along with the “7.2 User Acceptance Testing Plan Outline” examples in supporting attachments that CAMU will fully complete and upload into iMRMS as required.

CCAR model metrics and thresholds that are reviewed in pre and post implementation activities are discussed in detail in MDD section 9.2 below.

8. Operating and Control Environment

8.1 Access Control

[Provide evidences to show that the model will reside in a controlled and secured environment whereby access is controlled, and no un-authorized changes can be made to the model. Include persons with access to the model as of the date of this document in the table below.]

As stated above in MDD section 7.1, the environment in which the model will be implemented, executed, and backed up occurs on Bank’s centralized SAS Grid.

The following key personnel have read/write access to the proposed model, and this access is reviewed semi-annually as part of the global Employee Entitle Review System (EERS) review process.

Table 8.1 – Access Control Personnel

Name of Person with Access	Title	Group	Read or Read/Write Access	Reason for Access
Jenny Zhao	VP	SIMPLE	Read / Write Access	Responsible for conducting implementation testing and maintaining code with only read access
Jennifer Conner (or approved designate)	SVP	SIMPLE	Read Access	SIMPLE Team
Bin Sun (or approved designate)	Director	CAMU	Write Access	Model Development Lead
Jixiong Yu (or approved designate)	SVP	CAMU	Write Access	Production Lead
Gang Chen (approved designate: Michael Liu)	Director	CMI Loss Forecasting - Portfolio	Read Access	Model End User – Loss Forecasting
Jeff Armel	SVP	CMI Loss Forecasting - SFO	Read access	Model End User – Loss Forecasting / Risk Appetite Strategy
Bhootnath Singh (or approved designate)	Director	CPB Portfolio Risk	Read Access	Model End User – Loss Forecasting
Travis Spencer-Coye (or approved designate)	SVP	USRB Loss Forecasting	Read Access	Model End User – Loss Forecasting

8.2 Business Continuity

[Describe process and documentation in place to ensure continuity despite potential employee shifts and group shifts.]

Key process steps to ensure business continuity for the proposed model throughout its lifecycle are as follows:

- ❑ Comprehensive model documentation is completed throughout the entire life cycle including model development, independent validation, ongoing model input monitoring, quarterly performance assessments, and annual model review, up until model retirement.
- ❑ Comprehensive review/vetting occurs which includes In-business Maker-Checker process, peer review, Risk Senior Management review, Model End User/Sponsor review, MRM validation, Independent Risk Modeling Oversight validation, prior to model implementation and annual thereafter.
- ❑ Quarterly Ongoing Implementation Maintenance is completed which checks for operational root causes of poor model performance results arising from shifts in model input variables over time.
- ❑ Any change to the model will be reviewed and approved by the same process that is performed for new models. Depending on the materiality of the changes, either a Model Change Addendum or new MDD would be submitted for independent validation. Changes are implemented only after the required independent validation and approvals have occurred.
- ❑ There is a formal process for periodic review and approval for key modeling assumptions. The assumptions are reviewed annually by the key stakeholders during the model renewal process.
- ❑ Capacity is tracked in the Annual Plan which is submitted to Model Risk Management and Independent Risk Modeling Oversight and refreshed at a minimum every quarter thereafter.
- ❑ Staffing updates are reviewed monthly by Risk Management, and are part of the Model Sponsor Oversight Manager Control Assessment

8.3 Ongoing Quality Assurance process

[Describe the ongoing quality assurance (QA) process to monitor and ensure the operational reliability of the model.]

To ensure the ongoing quality assurance (QA) of the model input and output quality, the following processes and controls are performed:

1) Ongoing Model Input Monitoring (OMIM)

   Process: CAMU creates Ongoing Model Input Monitoring (OMIM) reports on internal and external data inputs to models based on frequency associated with model Materiality Risk Rating (MRR) in accordance with MRM Policy and GCCFRP Ch. 17 for all CCAR models. Key OMIM metrics include data quality review of potential variances in the population stability index and characteristic analysis as a result of shifts in model variables (dependent variable, independent variables, bureau variables, and macro-economic factors).

Control: CAMU reviews Ongoing Model Input Monitoring (OMIM) with model end users / sponsors at least quarterly to ensure that internal and external data inputs continue to be accurate, complete and consistent with model purpose and design. Significant variances in model inputs are escalated as applicable and discussed in the Model Risk Governance Committee (MGRC).

2) Ongoing Model Performance Assessments (OPA)

Process: CAMU generates accurate and timely Model MIS based on frequency associated with model Materiality Risk Rating (MRR) in accordance with MRM Policy and GCCFRP Ch. 17 for all CCAR models. Key OPA metrics include review of potential variances in cumulative error percentage (CERRPCT) for PD, EAD, LGD, GCL, NCL units/balance across short-term, medium-term, and long-term horizons.

Control: CAMU and relevant model stakeholders jointly review model MIS on a quarterly basis for applicable models to ensure performance metrics are within tolerance (or accompanied with appropriate rationale and/or compensating controls) and in compliance with MRM Policy and GCCFRP Chapter 17 requirements. MIS evidence is retained in iMRMS and on the shared directory. Significant variances in model inputs are escalated as applicable and discussed in the Model Risk Governance Committee (MGRC).

9. Ongoing Monitoring and Governance Plan

9.1 Ongoing Monitoring Plan

9.1.1 Monitoring Frequency

[Tick the following checkbox to confirm and agree to the monitoring frequency.]

☒ We (the Model Sponsor and Model Developer) understand and agree to comply with the following:

i. The required monitoring frequency depends on Model Risk Rating (MRR):
   - High: quarterly
   - Medium: semi-annually
   - Low: annually[2]

ii. The MRR will be assigned upon completion of model validation and is subject to change throughout the model lifecycle;

iii. We will conduct ongoing monitoring and submit monitoring reports to the validator based on the effective MRR in accordance of the required frequency.

Specifically for the proposed CCAR model, CAMU anticipates submitting monitoring reports to the validator(s) on a quarterly basis.

9.1.2 Monitoring Components

[Describe what components that will be included in the monitoring reports.]

The ongoing monitoring process to ensure quality assurance (QA) of the model inputs and outputs includes Ongoing Model Input Monitoring (OMIM) and Ongoing Model Performance Assessments (OPA) as described above in MDD section 8.3.

Using OPA as an example, the following steps are performed during the OPA process:

Stage 1: Analysis and Report (Performed by SIMPLE)

- ❑ Each quarter model performance monitoring results are generated by a team member (the Maker)
- ❑ A different team member (the Checker) verifies that the results are complete and accurate, i.e. by independently replicating the results and comparing them to the Maker's results
- ❑ The results are copied into the Quarterly Performance Monitoring (QPM) Word and Excel template by SIMPLE, and trigger breaches are noted within them
- ❑ The Checker ensures that the Summary QPM Word document and the Excel QPM Summary and Trend reflects the correct results and that the commentary is complete and accurate

[2] Annual Model Review (AMR) can be accepted in lieu of ongoing monitoring for low MRR models as long as all elements indicated in Section 9.1.2 are included in the AMR Sponsor document submitted by the Model Sponsor.

Stage 2: Stakeholder Feedback

- ❑ The Quarterly Performance Monitoring report is distributed by SIMPLE to stakeholders (Model Owners, Model Developers, etc.). Feedback is provided for all metrics which breach a trigger.

Stage 3: Publication

- ❑ SIMPLE then officially publishes the Quarterly Performance Monitoring Report to internal model stakeholders along with Model Risk Management (MRM) and Independent Risk Modeling Oversight (IRMO) for feedback.

The following specific OPA metrics are routinely analyzed as required by the loss forecasting performance testing guidance. These tests are focused on the key modelled variables, which are as follows:

- **GCL (AMT)** – For Unsecured portfolio, same as default exposure. In case of secured products, GCL is defined as Default Exposure adjusted with Loss Given Default.
- **NCL (AMT)** – the key tested variable in back-testing. NCL is defined as GCL adjusted with recoveries
- **GCL%ENR** –GCL%ENR is calculated on the quarterly level as a ratio of total newly defaulted exposure by the end of the quarter and the average non-defaulted exposure during the quarter for unsecured products. For secured products, GCL%ENR is adjusted with LGD.
- **NCL%ENR** – NCL%ENR is calculated on the quarterly level as a ratio of total NCL by the end of the quarter and the average non-defaulted exposure during the quarter

The following metrics are generated and reviewed as part of Quarterly Performance Assessments (QPA) for High Model Risk Level models:

- ❑ Quarterly Prediction Error Units (QERRunits)
- ❑ Quarterly Percent Prediction Error Units (QERRPCTunits)
- ❑ Cumulative Prediction Error Units (CERRunits)
- ❑ Cumulative Percent Prediction Error Balance (CERRPCTunits)
- ❑ Quarterly Prediction Error Balance (QERRbalance)
- ❑ Quarterly Percent Prediction Error Balance (QERRPCTbalance)
- ❑ Cumulative Prediction Error Balance (CERRbalance)
- ❑ Cumulative Percent Prediction Error Balance (CERRPCTbalance)
- ❑ Mean Absolute Deviation (MAD)
- ❑ Root Mean Squared Error (RMSE)
- ❑ Coefficient of Variation (COV)

The following metrics must also be updated in a global Model inventory quarterly for High Model Risk Level models:

- Cumulative Prediction Error Balance (CERRbalance)
- Cumulative Percent Prediction Error Balance (CERRPCTbalance)
- Mean Absolute Deviation (MAD)
- Root Mean Squared Error (RMSE)
- Coefficient of Variation (COV)

Back-testing thresholds have been established for Short term, Medium term and Long term and are applicable based on model usage, as defined in the following back-testing grid:

Table 9.1.2:- Back-testing Grid Loss Forecasting models

	Forecast Term		
	<= 12 months (<= 4 Qtrs)	13-39 months (5-13 Qtrs)	>39 months (14+ Qtrs)
Thresholds	CERRPCT - 20%	CERRPCT - 25%	CERRPCT - 40%

For loss forecasting, if CERRPCT on the rate variables at portfolio/segment/ component level breaches the above mentioned threshold, the model will be further assessed with respect to the Dual Matrix.

- The Quarterly and Cumulative Percent Prediction Errors are applicable to the overall Model and at the material segment level.
- Quarterly Percent Prediction Error must be reviewed to evaluate persistent bias in the time series. If a bias exists, then an evaluation must be done to correct the bias.
- Both unit and balance rates should be evaluated as outcomes may have different implications for component models.
- Possible adjustment actions include segment level calibrations, individual Model calibrations (slope and/or intercept adjustments), and overall Model adjustments (quantitative and/or qualitative).

Finally, CAMU works closely with Independent Risk Modeling Oversight to determine if there are any additional ongoing monitoring metrics specific to various portfolio nuances. This is typically determined post validation, during quarterly performance assessment reviews.

9.2 Assumptions Management Plan

[This section documents the process for periodic review and approval of key modeling assumptions. For statistically estimated models, this includes, among other things, the frequency of model coefficient re-estimation.]

There is a formal process for initial and periodic review and approval for key modeling and business assumptions. The assumptions noted in MDD section 5.3 above are reviewed annually by the key stakeholders during the annual model review process. Further, all model usage is attested to on a semi-annual basis by the appropriate GCB Business Head based on supporting evidence prepared by CAMU.

9.3 Model Usage Limitations Management

[This section describes how model usage limitations will be monitored and by whom. Include, for example, which reports of violated restrictions will be sent to whom; who will escalate restrictions violations to whom; and what actions will be taken on the violated restrictions.]

The proposed model will be used only for the approved purpose. Model end users / Sponsors will attest every quarter as part of ongoing model performance/QPM review as well as the semi-annual Model Attestation process that they are using the model as intended when the model was initially independently reviewed and approved. All limitations are tracked in iMRMS, each of which has a corresponding compensating control. Model usage limitations are monitored by CAMU and if violations of usage limitations are identified, this would be escalated within to the Model Governance Risk Committee (MGRC), Regional Modeling Head, and Executive Model Sponsor/Chief Risk Officer. Finally, MRM would be notified for guidance on next steps so that the model can be independently validated for any new usage.

9.4 Model Change Management Process and Approvals

[This plan describes the model change management process that must be followed for the given model, or refer to the change management plan for the broader business area (if applicable). All material changes to CCAR models must be reviewed and approved by the Model Risk Management. A log of approvals for changes to this model is provided below.]

Table 9.4 - Change Approvals Log

Version	Action	Proposer	Approver	Date Approved	Comment
1.0	9/30/18: Initial MDD Submission	CAMU			
1.1	11/21/18: Second MDD Submission (to address MEA feedback from MRM and IRMO)	CAMU			

9.5 Management Oversight

[Describe the process by which the upkeep and operation of the model is assessed by management. Include relevant reports and detail a list of process controls.]

The processes to assess model usage, continued use, model administration and governance, along with operational aspects of all NA Mortgage CCAR Models includes quarterly trigger reviews with Consumer Analytics Modeling Unit (CAMU), Strategic Initiatives, Model, & Project Lifecycle Execution (SIMPLE), and Global Model Risk Management (MRM) via detailed review of Quarterly Performance Assessment (QPA) reports. In addition, QPA reports are discussed and approved by committee stakeholders in the Model Governance Review Committee (MGRC), and also shared with MRM and Global Independent Risk. Implementation testing is performed on all CCAR models to ensure there is a secure version of the model production code (gold copy standard), to ensure SAS code warning/error messages are reviewed, well understood and if possible, rectified. NA Mortgage CCAR Model Ongoing Model Input Monitoring (OMIM) as well as Quarterly Performance Assessments (QPA) reviewed by CAMU Management and shared with Model Risk Management and Independent Risk Modeling Oversight throughout the model's lifecycle.

Further, all CCAR Models are validated by Independent Risk Modeling Oversight and discussed with IROC for functional soundness on an annual basis. The Regional Modeling Head (RMH) or designee reviews and approves the Inventory of mortgage risk models in each region, Model performance metrics, modeling personnel changes, and chairs the Model Governance Risk Committee (MGRC). The above key processes and related controls provide strong management oversight of the NA Mortgage CCAR Models which are performed by first and 2nd line of defense functions within the bank.

9.6 Model Governance

[Describe the model risk governance of the model. Transparently describe internal governance around the development of stress testing models and methodologies, and discuss how the stress testing methodologies have been implemented in the BHC's existing firm-wide risk management practices.]

Example:

The LookAhead Model was developed by the in-business Branded Cards Risk team. This was created with extensive consultations with Global Independent Risk, Global Modeling Oversight, in-business Branded Cards Risk and FP&A teams. The model proposal is reviewed by the Senior Country/Business Credit Manager (SCBCM) and Senior Scoring Specialist (SSS). After implementation, this model will be monitored as per the monitoring standards laid out by Bank Global Consumer Risk Management.

The proposed CCAR model was developed by the in-business Model Development Team known as Consumer Analytics Modeling Unit (CAMU), and internally checked by Strategic Initiatives, Model, & Project Lifecycle Execution (SIMPLE). These models were developed with extensive consultation with Model end users in Portfolio Risk Management, Senior Risk Management, Senior Management and Global Independent Risk. A series of internal checkpoints, coupled with presentations to the Independent Risk Operating Committee (IROC), occur throughout the Model development lifecycle to ensure strong models are developed in compliance with all applicable Bank policies. The NA Mortgage CCAR models are then independently reviewed for functional and technical soundness [collectively referred to as conceptual soundness] by Independent Risk and Model Risk Management respectively. In addition, Internal Audit (IA) completes an annual review of CCAR processes to review governance, data management, and reporting controls – as well as performs quarterly business monitoring activities.

Both the CAMU and SIMPLE teams have extensive experience in mortgage modeling, with several tenured staff, all of which hold advanced degrees in Statistics and Economics ranging between Masters and PHD proficiency levels. The monitoring standards and management oversight of the models are outlined in more detail in MDD sections 8.3, 9.2 and 9.5 above. The attached Committee Charters illustrate in detail the purpose, frequency, and stakeholder roles of the Model Governance Risk Committee (MGRC) and Independent Risk Oversight Committee (IROC).

Please refer to the "9.6_Model Governance Risk Committee Charter" along with the "9.6_Independent Risk Oversight Committee Charter" in supporting attachments.

10. Contingency Plan (for Vendor Models)

[This section is applicable to Vendor Models only.

Develop and document a contingency plan for vendor models in case the vendor is no longer available or capable of supporting the model. This contingency plan should be tailored based on the level of ongoing support required from the vendor and the business process in which the model is employed.]

This section is not applicable. The proposed model is not a vendor model.

11. References

[Please provide list of the references used to complete the MDD. Also attach documents if available]

1. Andrew Davidson (2006)-“Fixed-Rate Agency MBS Prepayments & Model Enhancements,” by Dan Szakallas, 2006.
2. Beyond Bond Inc. (2008); “Dynamic Econometric Loss Model – A default Study of US Subprime Markets “, Ted Hong, 2008
3. Biometrika (1984); Begg, C.B., Gray, R. (1984); “Calculation of polychotomous logistic regression parameters using individualized regressions.” 71, 11-18.
4. CoreLogic – CoreLogic® is the company financial services and real estate professionals turn to for comprehensive data, analytics and services
5. DeFranco, Ralph (2002) “Modeling Residential Mortgage Termination and Severity Using Loan Level Data”, Unpublished Doctoral Dissertation, Department of Economics, University of California, Berkley, 2002.
6. Edward N.C. Tong, Christophe Mues and Lyn Thomas (2013, “A zero-adjusted gamma model for mortgage loan loss given default” Volume 29, Issue 4, October–December 2013, Pages 548–562)
7. Foster, Chester, and Robert Van Order. 1984. “An Option-based Model of Mortgage Default.” Housing Finance Review, 3(4): 351–72.
8. LoanPerformance (2009) – LoanPerformance RiskModelTM Version 4; Technical Document, First American CoreLogic, 2009
9. Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta (2014) Mortgage modification and strategic behavior: evidence from a legal settlement with countrywide, The American Economic Review 104, 2830–2857.
10. Michael LaCour-Little and Yanan Zhang (2014, “Default Probability and Loss Given Default for Home Equity Loans,” Office of the Comptroller of the Currency Economics Working Paper 2014-1 June 2014)
11. Min Qi and Xiaolong Yang (2007, “Loss Given Default of High Loan-to-Value Residential Mortgages,” Office of the Comptroller of the Currency, OCC Economics Working Paper 2007-4 August 2007)
12. Quigley, J., and R. Van Order (1995): “Explicit Test of Contingent Claims Model of Mortgage Default”, Journal of Real Estate Finance and Economics, Volume 11
13. Yongheng Deng & John M. Quigley & Robert Van Order, 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options," Econometrica, Econometric Society, vol. 68(2), March.
14. Shirish Chinchalkar and Roger M. Stein, 2010. "*Comparing loan-level and pool-level mortgage portfolio analysis,*" Moody’s Research Labs, Current Draft: November 14, 2010

Appendix

15. Federal Reserve Board (2009), “The Supervisory Capital Assessment Program: Overview of Results,” May 7, 2009.
16. Federal Reserve Board (2009), “The Supervisory Capital Assessment Program: Design and Implementation,” April 24, 2009.
17. Federal Reserve Board (2009), “FAQs - Supervisory Capital Assessment Program.”
18. Federal Reserve Board (2011), “Comprehensive Capital Analysis and Review: Objectives and Overview,” March 18, 2011.
19. Federal Reserve Board (2012), “Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Predictions,” March 13, 2012.
20. Federal Reserve Board (2013), “Comprehensive Capital Analysis and Review 2013: Assessment Framework and Results,” March 2013.
21. Federal Reserve Board (2013), “Dodd-Frank Act Stress Test 2013: Supervisory Stress Test Methodology and Results,” March 2013.
22. Federal Reserve Board (2013), “2014 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule,” November 1, 2013.
23. Federal Reserve Board (2014), “Comprehensive Capital Analysis and Review 2014: Assessment Framework and Results,” March 2014.
24. Federal Reserve Board (2014), “Dodd-Frank Act Stress Test 2014: Supervisory Stress Test Methodology and Results,” March 2014.
25. Federal Reserve Board (2014), “2015 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule,” October 23, 2014.
26. Federal Reserve Board (2014), “Comprehensive Capital Analysis and Review 2015: Summary Instructions and Guidance,” October 2014.
27. Federal Reserve Board (2015), “Comprehensive Capital Analysis and Review 2015: Assessment Framework and Results,” March 2015.
28. Federal Reserve Board (2015), “Dodd-Frank Act Stress Test 2015: Supervisory Stress Test Methodology and Results,” March 2015
29. Federal Reserve Board (2016), “Dodd-Frank Act Stress Test 2016: Supervisory Stress Test Methodology and Results,” February 2016
30. Federal Reserve Board (2016), “Comprehensive Capital Analysis and Review 2016: Assessment Framework and Results,” July 2016.
31. Federal Reserve Board (2017), “Dodd-Frank Act Stress Test 2017: Supervisory Stress Test Methodology and Results,” June 2017
32. Federal Reserve Board (2017), “Comprehensive Capital Analysis and Review 2017: Assessment Framework and Results,” June 2017.

12. Appendix

[Please enclose any documents or supporting exhibits referenced in the template exhibits.]

Example: Scenario File, Vendor Supplier References, Grouping Results.

Abbreviation	Definition
ACAPS	Automated Credit Application Processing System
AFA	Alternative Forecasting Approach
AFT	Accelerated Failure Time, a type of survival analysis model
BHC	Bank Holding Company
BKFS	Black Knight Financial Services
BKFS TR	BKFS Technical Review
CAMU	Consumer Analytics Modeling Unit
CBSA	Core-Based Statistical Area
CCAR	Comprehensive Capital Analysis and Review
CCLTV	Current Combined Loan To Value ratio
CERR	Cumulative Error
CERRPCT	Cumulative Percent Prediction Errors
CFS	BankFinancial Services
CLTV	Combined Loan to Value ratio
CMI	BankMortgage Inc.
COV	Coefficient of Variation
CP	Cumulative Predicted outcome
CPB	Bank Private Bank
CPR	Conditional Prepayment Rate
CRO	Chief Risk Officer
CUSTOM	Bankcorp U.S. Total Online Management

Abbreviation	Definition
DFAST	Dodd Frank Annual Stress Testing
DFLT	Involuntary Prepayment; a term used by BKFS vendor Model
DPD	Days past due
DPLC	Default, Prepayment & Loss Curves
DRI	BankMortgage Non-performing loan servicing system
EAD	Exposure At Default
EL	Expected Loss
ENR	Ending Net Receivables
ERR	Error
ERRPCT	Error Percentage
FFIEC	Federal Financial Institutions Examination Council
FICO	Fair Isaac Co.
FRB	Federal Reserve Board
FRHEL	Fixed Rate Home Equity Loan
GCCFRP	Global Consumer Credit and Fraud Risk Policy https://policydirectory.Bank.net/cpd/_layouts/15/DocIdRedir.aspx?ID=CPDPROD-13-8866
GCL	Gross Credit Loss
GCRM	Global Country Risk Management (fka Bank Chief Economist Office)
GDP	Gross Domestic Product
GRMI	Global Risk Model Inventory
HARP	Home Affordable Refinance Program
HE	Home Equity
HELOC	Home Equity Line of Credit – open line of credit
HPI	House Price Index

Abbreviation	Definition
HTO	Housing Turnover
IA	Internal Audit
IR	Independent Risk
IROC	Independent Risk Operating Committee
IVP	Involuntary Pay-off
LGD	Loss Severity Given Default
LPM	Laser Pro Mortgage
LPS	Lender Processing Services (fka Black Knight Financial Services)
LTV	Loan To Value ratio
MA	Model Approach
MAD	Mean Absolute Deviation
MAP	Model Approval Package
MBA	Mortgage Business Association
MCA	Manager's Control Assessment
MDDP	Model Development Documentation Package
MDDT	Model Development Documentation Template
MGCV	Model Control, Governance and Validation
MGRC	Model Governance Risk Committee
MOB	Month on books
MOD	Modified loans
MRL	Model Risk Level
MRM	Model Risk Management
MRMC	Model Risk Management Committee
MSA	Metropolitan Statistical Area
MVG	Model Validation Group (=MRM)
NCL	Net Credit Loss
OCC	The Office of the Comptroller of the Currency

Abbreviation	Definition
OLTV	Original Loan To Value ratio
OTS	Office of Thrift Supervision
PCO	Partial Charge-off
PD	Probability of Default
PMACS	Portfolio Monitoring and Categorization System
RE	Real Estate
REFI	Refinance
REL	Bank’s U.S. Real Estate
REO	Real Estate Owned
RMH	Regional Modeling Head
RMSE	Root Mean Squared Error
S&S	Sold and Serviced loans
SAS	(the name of a statistical software)
SCBCM	Senior Country/Business Credit Manager
SR loans	Service-Released loans
SSS	Senior Scoring Specialist
UNIX	Interactive Time Sharing Operating System
UPB	Unpaid Principle Balance
USCCM	US Consumer, Commercial, and Mortgage
ZIP	Zone Improvement Plan; United States Postal Code

Template Version Control

Version	Change Description	Date	Updated by	Approved by
1.0	Initial template publication	07/03/2017	Yuri Yermakov Lonnie Cho	Steve Umlauf
1.1	Executive summary updated to include additional model change guidance	10/11/2017	Yuri Yermakov Jeanine Thompson	Steve Umlauf
2.0	Incorporated changes for updated Model Risk Management Policy	3/28/2017	Yuri Yermakov	Steve Umlauf
2.1	Updated to reflect changes in the revised Model Testing Guidance. Executive summary updated to cover all model usages	June 27, 2018	Yuri Yermakov	Steve Umlauf
2.2	Updated to align with the Document Automation Project: Added "Note on Document Automation Project" in the beginning of the document; deleted Model Info Sheet section; added instructions in Section 1 (Model Scope Purpose and Use), Section 2 (Limitations and Compensating Controls), Section 4 (Model Data) and Section 5 (Model Specifications) to indicate whether a sub-section would be completely replaced by a section within the automated document (in which case the user need not complete that sub-section) or that sub-section would serve as an additional information to the automated document	July 30, 2018	Yuri Yermakov, Jose Alcaraz	Steve Umlauf, Alan Kaplan

2.3	Separated The Model Testing Guidance document into 2 documents. (1) for PPNR & (1) for Retail Loss Forecasting. <u>Model Testing Guidance (MTG) for PPNR:</u> a. Guidance only contains information pertinent to PPNR models; b. GP's are the same as MTG for Loss Forecasting; c. Note on GP5 Segmentation has been added; d. Thresholds are now called "Recommended Thresholds"; e. In backtesting clarified on the use of Mean Absolute Deviation (MAD) and validator judgment; f. Other minor edits <u>MTG for Retail Loss Forecasting:</u> a. Guidance only contains information pertinent to Retail Loss Forecasting models; b. No change to the guidance (except for all PPNR information has been removed)	September 5, 2018	Yuri Yermakov	Steve Umlauf