# **Executive Summary**

This is a Validation report of the Wholesale PD Credit Risk Model (a.k.a., PD model) in accordance with the requirements of XX Model Risk Management Policy [3].

The objective of the Wholesale PD Credit Risk is to quantify the probability of default in the wholesale loan portfolio. Probability of default is used for appetite and capital adequacy testing purposes as well as for Business as Usual (“BAU”) stress testing for the XX and XX US Operations (XX). The model is utilized in credit risk limit setting and to project losses in the bank’s annual Risk Appetite Statement.

The PD model is used to forecast conditional default rates and rating transitions under macroeconomic scenarios developed internally and by XX, along with those provided by XX. The PDs are used to estimate losses and assess whether the BHC has sufficient capital to continue operations in times of financial stress. This model is applied to the Bank’s Wholesale portfolio, which primarily includes Commercial & Industrial (“C&I”) loans, leases, and letters of credit. The model is also used to evaluate the probability of a credit risk loss due to a trading counterparty failing to meet contractual obligations.

This report outlines the governance framework and model design of the rating transition model. It also presents the validation findings which cover model governance, input data, model design and methodology, implementation, the performance of the model, and documentation according to XX Model Risk Management policy.

# **Model Purpose and Use**

The Wholesale Credit Risk PD model is primarily intended to be used, in combination with the Loss Given Default (LGD), Exposure at Default (EAD), Counterparty Credit Risk (CCR), and balance projection models, to support loss forecasting for internal baseline projection and stress-testing under XX credit risk appetite and capital planning exercises. The model also supports ad hoc stress testing employed in the overall credit risk management framework.

The primary output of the model consists of forecasted PD estimates. The model supports forecasting horizons of any given length. For the purposes of the annual Risk Appetite statement, only the first four quarters are used.

The wholesale loan business is a traditional lending business whereby XX earns interest income for providing loans to institutional borrowers.

2.1 Model Overview

2.2 Development Data

Historical quarterly transition matrices were generated from XX Analytics EDF Database, using S&P rating data. The primary sources of independent variables include XX Analytics Data Buffet, the XX Reserve, and XX. Categories were devised for each of the 2,000+ distinct variables, and business line experts selected intuitive sets of variables from the categories to create the starting list of variables for each segment. Scenario forecasts were also provided by XX and the XX, but in the case of the Latin America variables, the Bank’s in-house economics department provided macroeconomic forecasts

2.3 Population Segmentation Schema

The models were segmented into 7 cohorts by industry type based on the portfolio composition of the entire MUSO portfolio: Energy, Auto, Financials, Utilities, Industrials, Commercial, and Latin America.

2.4 Modeling Approach Overview

The model consists of three parts: 1) The first part consists of a framework called ASRF (Asymptotic Single Risk Factor) that summarizes rating transition matrices using two parameters: m-factors, which represent systematic risk and correlations of the rating transition to systematic risk. 2) The second part consists of linkage models that establish a statistical relationship between m-factors and macroeconomic variables 3) The third part uses the results of the linkage model combined with the ASRF framework to calculate the forecasted rating transition matrix and ultimately the probability of default for XX portfolios.

An overview of the model structure and methodologies can be found in Error! Reference source not found... Please note that the model can be used to forecast the probability of default (which is the probability of transitioning into rating D), as well as the probability of transitioning into a different rating.

2.5 Model Structure

Asymptotic Single Risk Factor Model (ASRF)

The conditional transition matrix framework follows an industry standard default model (also commonly used in vended tools), which forms the basis for the standardized approach to credit risk adopted in the Basel II accords. Within this framework, a transition matrix (or subcomponent thereof) is described in terms of a single systemic factor (the “m factor”), representing how much an individual quarter’s rating transitions deviate from the average historical experience. The m-factor, which takes the form of a draw from a standard normal distribution, is positive under favorable conditions and negative under adverse conditions, with the magnitude corresponding to the distance from the average matrix. This is a structural Merton-style model, an obligor is said to be in default if the value of the total assets, V, drops below a certain threshold e.g. the contractual value of its obligations. The probability of default, thus, is given by

𝑃(𝑌=1)=𝑃(𝑉<𝑐𝑖)

In the so-called one-factor models, defaults and rating transitions are driven by a single factor, a continuous, normally distributed underlying credit indicator, X, which can be further decomposed into a systemic component, M, shared by all borrowers and an idiosyncratic, borrower specific component, X, Z, and ε, which are related through the correlation parameter ρ as follows, where M and Y are independent unit normally distributed random variables: are standard normal variables and mutually independent, the default probability may then be expressed as the probability of a standard normal variable falling below a critical value, defined with respect to the different ratings (with a total of n rating classes or risk rating grades).

The probability of default can now be formulated as

𝑃(𝑌𝑖,𝑡=1)=𝑃(𝑋𝑖,𝑡<𝑐𝑖)=𝛷(𝑐𝑖)

 Since

𝑋𝑖,𝑡∼𝑁(0,1)

it follows that

𝑃(𝑋𝑖,𝑡<𝑐𝑖)=𝛷(𝑐𝑖)⇔𝛷−1(𝑋𝑖,𝑡)=𝑐𝑖

 This is the unconditional default probability. If the outcome of the systematic risk factor is known, we could calculate the conditional probability of default

𝑃(𝑌𝑖,𝑡=1∣∣𝑀𝑡=𝑚)=𝑃(𝑋𝑖,𝑡<𝑐𝑖∣∣𝑀𝑡=𝑚) =𝑃(√𝜌𝑀𝑡+√1−𝜌𝜀𝑖,𝑡<𝑐𝑖∣∣𝑀𝑡=𝑚) =𝑃(𝜀𝑖,𝑡<𝑐𝑖−√𝜌𝑀𝑡√1−𝜌∣∣𝑀𝑡=𝑧) =𝛷(𝑐𝑖−√𝜌𝑚√1−𝜌)

 The correlation parameter, which determines the extent to which transition behavior is driven by systemic vs. idiosyncratic factors, is derived empirically, so that the distribution of observed m-factors exhibits unit variance. Given an average transition matrix and correlation parameter ρ, a conditional transition matrix can then be expressed in terms of the m-factor.

It is assumed that, conditional on an initial rating i at the beginning of a quarter, one partitions values of the credit change indicator X into a set of disjoint bins. Conditional on an initial credit rating (G) at the beginning of a quarter, the cumulative probability distribution of the credit change indicator X can then be partitioned into a series of discrete bins(xgG,xg+1G), where g represents the sequence of PD ratings.

The figure below illustrates the relationship between the discrete bins and the observed transition frequencies for a given rating grade (risk rating 5).In the conditional transition matrix framework, the probability of transitioning from rating grade G to g, conditional on systemic factors M in period “t” is:

𝑃(𝐺,𝑔∣∣𝑀𝑡)=𝛷(𝑥𝑔+1𝐺−√𝜌𝑀𝑡√1−𝜌)−𝛷(𝑥𝑔𝐺−√𝜌𝑀𝑡√1−𝜌)

 Within this framework, empirical m-factors (Mt) are derived for historical matrices to minimize the difference between the modeled transition rates and the actual transition rates for a given period. An illustration of this is shown in Figure 2 below. Under adverse economic conditions, the normal distribution of rating migration would shift to the right, implying worse ratings levels, meaning that the probability of downgrade and default increases. As the credit migration matrices are driven by a single parameter M, which depicts the average financial health of firms, this shift corresponds to a simple change in the value of M. Once the quarterly and average migration matrices are created, an m-factor can be calculated on a quarterly basis such that for a fixed correlation (ρ), the reconstructed migration matrix (based on that m-factor, the average migration matrix, and the correlation assumption) is as “close” as possible to the actual matrix for that quarter. The asset correlation parameter was estimated separately so that the resulting time series of m-factors has a unit variance (Forest, 1998). To do so, an initial correlation parameter was first assumed and the resulting empirical m-factors were estimated. A numerical search algorithm was then used to fit the correlation parameter based on iteratively repeating the process for higher asset correlations until the chosen parameter returned a series of m-factors with a variance close to 1.

Linkage model between M-Factors and Macro variables

M-factors are regressed vs. macro factors and a linkage model is selected. The OLS model takes the following form:

The factors were selected based on a process of careful factor selection based on business intuition and then by applying single factor regression in order to determine which macro variables had a strong statistical relationship to the m-factors.

# **Independent Variables**

Bank has subscribed to Moody’s Data Buffet to obtain macroeconomic indicators for the model development. These indicators represent important economic sectors of the UAE and reflect the economic outlook of the country in general.

## **Moody’s Scenario Narratives**

Data Buffet provides both historical and forecasted data for a set of macroeconomic indicators in its inventory. Data is available on quarterly frequency for multiple economic scenarios that are built considering various factors/ narrations that impact both regional and global economic outlook. For some indicators, historical data is available from as early as Q2-1952. When the data was downloaded for model development, Data Buffet contained forecasts from Q4-2024 to Q4-2053. Since bank uses three scenarios – upside, baseline and downside – for ECL calculation, forecasts for these scenarios were selected for calculating macroeconomic overlay. Below is the summary of key assumptions of Moody’s to derive/ construction of these scenarios as on Sep-2024.

1. **Baseline Scenario Forecast**

This scenario is the baseline forecast of Moody’s Analytics. Since it is baseline, probability that the economy will perform better than this projection is equal to 50%, the same as the probability that it will perform worse.

1. **Upside Scenario Forecast**

In this scenario the economy will outperform the baseline. Probability that economy will outperform this scenario is 10% and probability that the economy will underperform this scenario is 90%.

1. **Downside Scenario Forecast**

The economy underperforming the baseline. The probability that the economy will perform worse than in this scenario is 10%. The probability that it will perform better than in this scenario is 90%.

## **Macroeconomic Variables**

Moody’s analytics updates the forecasts on quarterly basis, considering all the events that occurred in the completed quarter, which may have effect on the macroeconomic indicators. Forecasts obtained for the model development were updated in Sep-2024. Following UAE specific macroeconomic indicators were selected from the Moody’s Data Buffet, to be used as independent variables for model development:

1. Real Gross Domestic Product [GDP], (Bil. 2010 AED)
2. Gross domestic product [GDP] - Real - Abu Dhabi, (Mil. 2017 USD)
3. Gross domestic product [GDP] - Real - Dubai, (Mil. 2017 USD)
4. Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP], (Bil. 2017 Intl. USD)
5. Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP] - Per capita, (Ths. 2017 Intl. USD)
6. General government balance to GDP ratio, (%)
7. General government debt to GDP ratio, (%)
8. General Government Gross Debt, (Bil. AED)
9. House Price Index: Real, (Index 2010=100)
10. House Prices: Residential properties - Abu Dhabi and Dubai, (Index 2010=100)
11. Residential property prices - Abu Dhabi: All dwellings, (Index 2010=100)
12. Residential property prices - Dubai: All dwellings, (Index 2010=100)
13. Residential property prices - Abu Dhabi: All dwellings, (AED per m²)
14. Residential property prices - Dubai: All dwellings, (AED per m²)
15. Construction Permits: Residential buildings - Abu Dhabi, (Ths. #, N)
16. Hotel occupancy - Abu Dhabi, (%)
17. Hotel occupancy - Dubai, (%)
18. Energy Production - Crude oil including lease condete, (Mil. Bbl per day)
19. Commodity prices: Crude oil [Dubai Fateh], (USD per Bbl)
20. Consumer Price Index, (Index 2021=100)
21. Interest Rate: 3-month EIBOR, (% p.a.)
22. Share Price Index: ADX General Index, (Index)
23. Labor Force Survey: Unemployment Rate, (%)
24. Unit labour cost, (Index 2010=100)
25. Compensation of Employees - Real, (Bil. 2010 AED)
26. General Government Finance: Revenue, (Bil. AED)
27. Real Domestic Demand, (Bil. 2010 AED)
28. Real Exports of Goods and Services, (Bil. 2010 AED)

## **Variable Transformation**

All the above variables were transformed to facilitate efficient and accurate assessment of nature and extent of their relationship with the dependent variable (default rate). Following transformations were applied on each macroeconomic variable:

Transformations Applied on Independent Variables

|  |  |  |
| --- | --- | --- |
| **Transformation** | **Formula** | **# Variables Created** |
| One quarter percentage change |  | 5 variables.  Transformed variable and four lags of it. |
| Two quarter percentage change |  | 5 variables.  Transformed variable and four lags of it. |
| Three quarter percentage change |  | 5 variables.  Transformed variable and four lags of it. |
| Four quarter percentage change |  | 5 variables.  Transformed variable and four lags of it. |
| Logged one quarter change |  | 5 variables.  Transformed variable and four lags of it. |
| Logged two quarter change |  | 5 variables.  Transformed variable and four lags of it. |
| Logged three quarter change |  | 5 variables.  Transformed variable and four lags of it. |
| Logged four quarter change |  | 5 variables.  Transformed variable and four lags of it. |
| 2-Qaurter moving average |  | One variable |
| 3-Qaurter moving average |  | One variable |
| 4-Qaurter moving average |  | One variable |
| Lags of original variable | * One quarter lag * Two quarter lag * Three quarter lag * Four quarter lag | Four variables |
| Natural log of original variable |  | 5 variables.  Transformed variable and four lags of it. |
| Standardization | Where,  = average of IV  = standard deviation of IV | 5 variables.  Transformed variable and four lags of it. |
| Min-Max Scaling |  | 5 variables.  Transformed variable and four lags of it. |
| Mean Normalization |  | 5 variables.  Transformed variable and four lags of it. |
| Square root Transformation |  | 5 variables.  Transformed variable and four lags of it. |
| Centre Mean Transformation | Where,  = average of MEV | 5 variables.  Transformed variable and four lags of it. |

In total, 2079 independent variables were created by applying the above transformations.

## **Variable Reduction**

As there were 2079 independent variables, developing regression models for all the combination of these variables will require huge time and computational effort. Additionally, not all the variables will have sufficient causal relationship with the dependent variable. Hence, only those transformations that had optimal predictive power and economic meaning of the original variables, with reference to dependent variable, were selected for the model exploration process.

This section describes the methods used to reduce the number of independent variables to select variables with good predictive power.

## **Intuitive Relationship with Dependent Variable**

Table below illustrates the a-priori/ expected relationship of each independent variable with the dependent variable. Default rates are expected to be higher in a downside economic scenario and compared to upside economic scenario. Forecasts of these selected variables, under all three scenarios, were plotted on graphs/ charts to determine the expected relationship with the default rate.

For example, Gross Domestic Product [GDP] is expected to be higher in upside scenario than baseline and adverse scenarios. And hence, it has inverse relationship with the default rate. However, forecasts for the Unemployment Rate are expected to be higher in downside scenario compared to baseline and upside scenarios. And hence, it has direct relationship with the default rates.

It is expected that corresponding transformations of a macroeconomic variable must have same relationship with default rate as that of original variable.

Expected Relationship of Independent Variables with Dependent Variable

|  |  |
| --- | --- |
| Independent Variable | Expected Relationship |
| Real Gross Domestic Product [GDP], (Bil. 2010 AED) | Inverse |
| Gross domestic product [GDP] - Real - Abu Dhabi, (Mil. 2017 USD) | Inverse |
| Gross domestic product [GDP] - Real - Dubai, (Mil. 2017 USD) | Inverse |
| Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP], (Bil. 2017 Intl. USD) | Inverse |
| Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP] - Per capita, (Ths. 2017 Intl. USD) | Inverse |
| General government balance to GDP ratio, (%) | Inverse |
| General government debt to GDP ratio, (%) | Direct |
| General Government Gross Debt, (Bil. AED) | Direct |
| House Price Index: Real, (Index 2010=100) | Inverse |
| House Prices: Residential properties - Abu Dhabi and Dubai, (Index 2010=100) | Inverse |
| Residential property prices - Abu Dhabi : All dwellings, (Index 2010=100) | Inverse |
| Residential property prices - Dubai: All dwellings, (Index 2010=100) | Inverse |
| Residential property prices - Abu Dhabi: All dwellings, (AED per m²) | Inverse |
| Residential property prices - Dubai: All dwellings, (AED per m²) | Inverse |
| Construction Permits: Residential buildings - Abu Dhabi, (Ths. #) | Inverse |
| Hotel occupancy - Abu Dhabi, (%) | Inverse |
| Hotel occupancy - Dubai, (%) | Inverse |
| Energy: Production - Crude oil including lease condensate, (Mil. Bbl per day) | Inverse |
| Commodity prices: Crude oil [Dubai Fateh], (USD per Bbl) | Inverse |
| Consumer Price Index, (Index 2021=100) | Inverse |
| Interest Rate: 3-month EIBOR, (% p.a.) | Inverse |
| Share Price Index: ADX General Index, (Index) | Inverse |
| Labor Force Survey: Unemployment Rate, (%) | Direct |
| Unit labor cost, (Index 2010=100) | Inverse |
| Compensation of Employees - Real, (Bil. 2010 AED) | Inverse |
| Real Domestic Demand, (Bil. 2010 AED) | Inverse |
| Real Exports of Goods and Services, (Bil. 2010 AED) | Inverse |

## **Expected Predictive Power**

Pearson’s correlation coefficient was computed for each transformation to assess the nature of relationship with the dependent variable. A positive correlation coefficient is expected for an independent variable that has direct relationship with the dependent variable. A negative correlation coefficient is expected for an independent variable that has inverse relationship with the dependent variable. All the variables that had counterintuitive relationship with the dependent variable were excluded during the preliminary variable selection process.

There were 1,408 transformations that had a-priori business intuitive relationship with the dependent variable. Developing models based on all combinations of these many transformations will require huge resources and time. Hence, these variables were further reduced based on their expected predictive power. While business intuition of an independent variable was assessed based on sign of the independent variable’s correlation coefficient, its magnitude was used to assess a variable’s expected predictive power.

Different independent variables (transformations) will have varying levels of causal relationship with the dependent variable. Additionally, number of intuitive transformations selected for each variable in the previous step varies significantly. Hence, varying correlation coefficient cut-offs were applied on each independent variable to shortlist:

* reasonable number of transformations under each variable, and
* transformations that have optimum predictive power

Table below illustrates the correlation threshold used for each variable and number of transformations selected.

Macroeconomic Variable Transformations with Intuitive Relationship

|  |  |  |
| --- | --- | --- |
| Independent Variable | Correlation Threshold | # Transformations Selected |
| Real Gross Domestic Product [GDP], (Bil. 2010 AED) | 0.43 | 10 |
| Gross domestic product [GDP] - Real - Abu Dhabi, (Mil. 2017 USD) | 0.36 | 10 |
| Gross domestic product [GDP] - Real - Dubai, (Mil. 2017 USD) | 0.25 | 8 |
| Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP], (Bil. 2017 Intl. USD) | 0.44 | 8 |
| Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP] - Per capita, (Ths. 2017 Intl. USD) | 0.44 | 6 |
| General government balance to GDP ratio, (%) | 0.35 | 5 |
| General government debt to GDP ratio, (%) | 0.53 | 6 |
| General Government Gross Debt, (Bil. AED) | 0.57 | 4 |
| House Price Index: Real, (Index 2010=100) | 0.57 | 9 |
| House Prices: Residential properties - Abu Dhabi and Dubai, (Index 2010=100) | 0.63 | 9 |
| Residential property prices - Abu Dhabi: All dwellings, (Index 2010=100) | 0.52 | 10 |
| Residential property prices - Dubai: All dwellings, (Index 2010=100) | 0.655 | 10 |
| Residential property prices - Abu Dhabi: All dwellings, (AED per m²) | 0.52 | 10 |
| Residential property prices - Dubai: All dwellings, (AED per m²) | 0.655 | 10 |
| Construction Permits: Residential buildings - Abu Dhabi, (Ths. #) | 0.25 | 4 |
| Hotel occupancy - Abu Dhabi, (%) | 0.25 | 2 |
| Hotel occupancy - Dubai, (%) | 0.47 | 9 |
| Energy: Production - Crude oil including lease condensate, (Mil. Bbl per day) | 0.30 | 12 |
| Commodity prices: Crude oil [Dubai Fateh], (USD per Bbl) | 0.67 | 13 |
| Consumer Price Index, (Index 2021=100) | 0.30 | 6 |
| Interest Rate: 3-month EIBOR, (% p.a.) | 0.35 | 6 |
| Share Price Index: ADX General Index, (Index) | 0.30 | 6 |
| Labor Force Survey: Unemployment Rate, (%) | 0.34 | 10 |
| Unit labor cost, (Index 2010=100) | 0.33 | 8 |
| Compensation of Employees - Real, (Bil. 2010 AED) | 0.39 | 10 |
| Real Domestic Demand, (Bil. 2010 AED) | 0.25 | 2 |
| Real Exports of Goods and Services, (Bil. 2010 AED) | 0.55 | 8 |
| **Total** |  | **211** |

Regression models were developed by regressing the dependent variable (logit transformed observed default rate series) on multiple combinations made of three unique macroeconomic variables from the above 211 transformations. More detailed methodology has been provided in the next section.

# **Model Development**

## **Linear Regression**

Influence of macroeconomic variables on the portfolio default rate was estimated using linear regression model. Regression analysis is a statistical technique that can be used to analyse the relationship between a single dependent variable and one or more independent variables. The objective of the regression analysis is to use the independent variables, whose values are known, to predict the values of dependent variable. There are several uses of regression analysis:

* To predict future economic conditions, trends, or values
* To determine the relationship between two or more variables
* To understand how one variable change when another change

A regression model that establishes the relationship between a dependent variable and an independent variable is called a simple regression. A regression model that uses several explanatory variables to predict the outcome of a response variable is called multiple linear regression (MLR), also known simply as multiple regression. Multiple regression analysis takes the following form:

Where, for i = n observations:

= dependent variable

= explanatory variables

= intercept (constant term)

​= slope/ regression coefficients for each explanatory variable

= error term (also known as the residuals)

​Ordinary Least Squares (OLS) is the most common estimation method for estimating the linear models. Regression analysis, an inferential technique, is built on OLS procedure. Goal of the regression analysis is to draw a random sample from a population and use it to estimate the population parameters – coefficients are the estimates of population parameters. In estimating these parameters, OLS procedure makes some assumptions. Following are the important assumptions of OLS regression:

* There is a linear relationship between the dependent variable and the independent variables.
* The error term is normally distributed with zero mean and constant variance
* Observations of the error term are uncorrelated with each other (no autocorrelation)
* No independent variable is a perfect linear function of other explanatory variables (no multicollinearity)

The coefficient of determination (R-squared) is the most commonly used measure of predictive accuracy for the regression model. It measures how much of the variation in dependent variable that can be explained by the variation in the independent variables. Values of R2 range from 0 to 100. Where zero R2 indicates no relationship between dependent and independent variables; 100 indicates all the variation in the dependent variable is explained by the independent variables included in the regression model.

However, the R-squared increases for every additional independent variable in the model, even if due to chance alone. It never decreases. Consequently, a model with more independent variables may appear to have a better fit simply because it has more independent variables.

On the other hand, the adjusted R-squared takes into account the number of independent variables used for predicting the dependent variable. In doing so, one can determine whether adding new variables to the model actually increases the model fit.

Logit transformed quarterly default rate series was used as dependent variable to develop macroeconomic model for the personal loan portfolio.

## **Number of Variables in the Model**

The sample size used in the regression analysis has a direct impact on the appropriateness and statistical power of the regression model. Final dataset used for the model development contained 47 observations from Q1-2012 to Q3-2023.

To enhance the model generalizability and to maintain required observations-to-variables ratio, models that contain three independent variables were developed.

## **Variable Combinations**

The final dataset contained 211 independent variables/ transformations, that were shortlisted based on the variable selection criteria described in the previous section. All the possible combinations of these independent variables, with three independent variables in each combination, were created.

Where,

= total number of independent variables.

= number of independent variables to be selected from the total independent variables, which is 3 in this context.

A total of 1,543,465 3-variable combinations were created for developing the model using 211 shortlisted independent variables. The variable combinations that met the following criteria were excluded from model development process:

* Combinations that contained transformations of the same variable repeating more than once.
* Combinations containing the variables that measure identical macroeconomic event. The table below shows the group of variables/ indicators that measure same/ similar macroeconomic event.

|  |
| --- |
| * Real Gross Domestic Product [GDP], (Bil. 2010 AED) * Gross domestic product [GDP] - Real - Abu Dhabi, (Mil. 2017 USD) * Gross domestic product [GDP] - Real - Dubai, (Mil. 2017 USD) * Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP], (Bil. 2017 Intl. USD) * Real Gross Domestic Product [GDP] - Purchasing Power Parity [PPP] - Per capita, (Ths. 2017 Intl. USD) |
| * General government debt to GDP ratio, (%) * General Government Gross Debt, (Bil. AED) |
| * House Price Index: Real, (Index 2010=100) * House Prices: Residential properties - Abu Dhabi and Dubai, (Index 2010=100) * Residential property prices - Abu Dhabi: All dwellings, (Index 2010=100) * Residential property prices - Dubai: All dwellings, (Index 2010=100) * Residential property prices - Abu Dhabi: All dwellings, (AED per m²) * Residential property prices - Dubai: All dwellings, (AED per m²) |
| * Hotel occupancy - Abu Dhabi, (%) * Hotel occupancy - Dubai, (%) |
| * Real Imports of Goods and Services, (Bil. 2010 AED) * Real Exports of Goods and Services, (Bil. 2010 AED) |

About 1,010,677 unique variable combinations were left for developing regression models, after applying the above exclusion criteria.

## **Model Development**

Total of 1,010,677 regression models were developed by using multiple regression analysis technique for the personal loan portfolio.

## **Model Selection Criteria**

Of all the 1,010,677 models developed, only those models that met following requirement were selected for further analysis.

**Model Selection Criteria**

|  |  |
| --- | --- |
| Criteria | Remarks |
| Business intuition | * Regression coefficients must have business intuitive sign. * Variables included in the model should be relevant to the portfolio. |
| Significant coefficient | If the regression coefficient is significant at the 0.05 level, then null hypothesis is rejected and accept the alternative hypothesis that a relationship exists between the dependent and independent variable(s). |
| Model significance | F-test is used to test the overall significance of a regression model. A significant F-test (p<0.05) provide sufficient evidence to conclude that the regression model fits the data better than the model with no independent variables. |
| Adjusted R-square | Models with 60% and above adjusted R-square were selected for the further analysis. |
| OLS Assumptions | Selected model(s) must meet all the classical assumptions of OLS regression models:   * Normality * Linearity * Homoscedasticity * Autocorrelation |
|  | A stationary process has the property that the mean, variance and autocorrelation structure do not change over time. The Augmented Dicky-Fuller test (ADF test) was used to test whether the residuals are stationary. If the p-value of the ADF test is <0.05 then residuals are stationary. |
| Intuitiveness of forecasts | It is expected that default rates are higher in the downside economic scenario compared to baseline scenario, and lower in upside scenario compared to baseline scenario.  Models that predicted counterintuitive forecasts were discarded. |

Technical details on all the above-mentioned statistical tests can be found in appendix at the end of this document

We found only 350 regression models met all the above statistical requirements, including the economically intuitive regression coefficients. These models were further evaluated to assess whether they produce intuitive forecasts of default rates under different economic scenarios. This was done through plotting the forecasts on a plot. It is expected that, selected model must predict a higher default rate for downside scenario, compared to baseline & upside scenario, and lower default rate for upside scenario, compared to baseline & downside scenarios.

# **Selected Model**

Though many models met the selection criteria listed in the previous section, two preferred models were selected for the personal loan portfolio. In addition to statistical criteria discussed in the previous section, models were selected considering relevance of macroeconomic indicator in the selected model, with respect to portfolio for which the model is being developed. Table below illustrates the first preferred model, which will be used to derive the IFRS9 compliant forward-looking point-in-time (PiT) PD term structure. Please refer to appendix-1 for the derails on other preferred models.

## **Model Statistics**

Basic Statistics for the Selected Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Macroeconomic Variable | Transformation | Coefficients | Std. Error | P-value |
| Intercept | - | -3.5944 | 0.2432 | 0.0000 |
| General government debt to GDP ratio, (%) | Log Transformation | 0.2601 | 0.0762 | 0.0014 |
| House Price Index: Real, (Index 2010=100) | Logged four quarter change – Lag1 | -1.9064 | 0.2552 | 0.0000 |
| Energy Production - Crude oil including lease condensate, (Mil. Bbl per day) | Four quarter percentage change – Lag4 | -2.3218 | 0.2846 | 0.0000 |

|  |  |
| --- | --- |
| F Stat | 52.27 |
| Adjusted R Square | 78.28 |
| Standard Error | 15.01 |

It is evident from the above table that coefficients of the macroeconomic variables included in the model are business intuitive and overall model is significant. Model has an adjusted r-square of 78.28%, which indicates a fair fit and predictive power.

Chart below demonstrates the observed default rates as predicted by the selected model in the sample used for the model development. It is observed that predicted default rate series closely follows the observed default rate series in the model development sample.

Observed & Predicted Default Rates in Development Period

## **Testing for OLS assumptions**

Standard tests that are widely used in the industry were used to assess whether the selected model comply with the OLS assumptions. Table below illustrates the tests and the outcome of the tests.

Tests for OLS Assumptions

|  |  |  |  |
| --- | --- | --- | --- |
| **Assumption** | **Test** | **Test Statistic** | **p-value** |
| Normality | Jarque-Bera Test | 1.0938 | 0.5787 |
| Shapiro-Wilk | 0.9821 | 0.6794 |
| Anderson-Darling | 0.3068 | >0.050 |
| Linearity | Rainbow Test | 0.4532 | 0.9658 |
| Homoscedasticity | White | 12.2000 | 0.2023 |
| Breusch–Pagan test | 5.7751 | 0.1231 |
| Autocorrelation | Durbin-Watson | 1.2214 | Significant @ 1% level |
| Multicollinearity | Variance Inflation Factor (VIF) | 1.3101 | Maximum VIF of all 3 variables |

It is evident from the above table that the selected model meets all the underlying assumptions of OLS procedure, suggesting the model is an unbiased, linear and best estimator of the personal loan portfolio default rates. Please refer to the appendix for the detailed description of the above tests.

**Test for Autocorrelation**

To infer whether the errors are auto-correlated, calculated Durbin-Watson (DW) statistic is compared with lower (DL) and upper (DU) bounds in the Durbin-Watson table, at 1% significance level. Findings from this comparison are interpreted as follows:

* if DW < DL: Positive correlation exists.
* if (4-DW) < DL: Negative correlation exists.
* if DW > DU: No Correlation
* if (4-DW) > DU: No Correlation
* If the DW or (4-DW) is between two bounds, the test is inconclusive.

Null hypothesis of zero autocorrelation in the residuals against the alternative hypothesis that residuals are not autocorrelated was tested. As there are 47 observations in the model, corresponding lower (DL) and upper (DU) bounds for n=45 are 1.201 and 1.474, respectively. By comparing the test statistics with the bounds, it can be determined that, there is no evidence of positive or negative correlation.

In addition to Durbin-Watson test, model was further checked to evaluate any possible effects of autocorrelation using Newey-West estimator. Regression coefficients remained significant post adjusting the standard errors using Newey-West test, indicating that the model is heteroskedasticity and autocorrelation consistent.

Newey-West Adjusted Standard Errors

|  |  |  |
| --- | --- | --- |
| Variable | Std. Error | P-value |
| Intercept | 0.3246 | 0.0000 |
| General government debt to GDP ratio, (%) | 0.0971 | 0.0074 |
| House Price Index: Real, (Index 2010=100) | 0.4027 | 0.0000 |
| Energy Production - Crude oil including lease condensate, (Mil. Bbl per day) | 0.2480 | 0.0000 |

## **Test for Stationarity**

Residuals must be stationary to avoid spurious regressions. The Augmented Dicky-Fuller test (ADF test) was used to test whether the residuals are stationary. A Lag equivalent to 12\*(N / 100) {1/4} was used to perform the test, where N is the number of observations. For this model, this translates to 10 lags.

|  |  |
| --- | --- |
| ADF test statistic | -4.4448 |
| p-value | 0.0002 |

The p-value for the ADF test statistic is <0.05 indicates that the variables included in the model development are stationary and hence model predictions are stable and reliable.

## **Model Variable Forecasts**

Charts below shows the Moody’s forecasts for the macroeconomic variables in the selected model under three economic scenarios.

General government debt to GDP ratio, (%)

House Price Index: Real, (Index 2010=100)

Energy Production - Crude oil including lease condensate, (Mil. Bbl per day)

It is evident from the above charts of macroeconomic variables that adverse scenario forecasts are clearly well separated from their corresponding baseline forecasts. However, upside scenario forecasts are much closer to the baseline forecasts, especially in Energy Production variable followed by House Price Index variable. It reflects the Moody’s scenario narratives discussed earlier under section 2 that upside scenarios are not too optimistic compared baseline scenario given the challenging economic and geopolitical situations. Hence, it is expected that predicted default rates using these macroeconomic forecasts would result in default rates for upside scenario that is not too different from the baseline scenario.

## **Intuitiveness of Forecasts**

Selected model will be used to predict the default rates for three hypothetical economic scenarios – upside, baseline and downside. It is expected that, selected model must predict a higher default rate for the downside scenario, compared to baseline and upside scenarios.

Forecasted Default Rates Under Different Scenarios

Given that upside scenario forecasts of the macroeconomic variables are not significantly different from baseline forecasts, resulting default rate predictions also follow the same trend. However, it can be observed that upside default rate predictions do not cross-over the baseline predictions, though they look converging into baseline predictions. Table below provides the default rate predictions for all three scenarios to illustrate that upside scenario predictions are lower than the baseline scenario predictions.

|  |  |  |  |
| --- | --- | --- | --- |
| Quarter | Upside | Baseline | Adverse |
| 2025Q1 | 5.26% | 5.47% | 6.14% |
| 2025Q2 | 5.04% | 5.39% | 6.76% |
| 2025Q3 | 5.01% | 5.48% | 7.57% |
| 2025Q4 | 4.96% | 5.59% | 8.47% |
| 2026Q1 | 5.17% | 5.87% | 8.82% |
| 2026Q2 | 4.97% | 5.61% | 8.13% |
| 2026Q3 | 4.78% | 5.34% | 7.41% |
| 2026Q4 | 4.66% | 5.10% | 6.69% |
| 2027Q1 | 4.59% | 4.91% | 6.07% |
| 2027Q2 | 4.68% | 4.93% | 5.71% |
| 2027Q3 | 4.72% | 4.93% | 5.43% |
| 2027Q4 | 4.76% | 4.95% | 5.23% |
| 2028Q1 | 4.79% | 4.94% | 5.04% |
| 2028Q2 | 4.81% | 4.93% | 5.00% |
| 2028Q3 | 4.83% | 4.94% | 5.00% |
| 2028Q4 | 4.86% | 4.97% | 5.03% |
| 2029Q1 | 4.89% | 5.01% | 5.09% |
| 2029Q2 | 4.92% | 5.04% | 5.15% |
| 2029Q3 | 4.94% | 5.07% | 5.19% |
| 2029Q4 | 4.97% | 5.10% | 5.25% |

Default rates forecasted from this model will be used to apply the macroeconomic overlay on the PD estimates of personal loan portfolio, to estimate forward-looking ECL. This is achieved by process called PD calibration. Step-by-step detail of this process is documented in the Retail portfolio ECL calculation document.