5.2 PD model stress test

5.2.1 Kolmogorov–Smirnov Test

The KS statistic is used to measure the discriminatory power of the PD model. It is defined as the maximum difference between the cumulative percentage of good samples (i.e., non-defaulters) and the cumulative percentage of bad samples (i.e., defaulters). A higher KS value implies a good fit of the model.

The test returns a KS value.

|  |  |
| --- | --- |
| Ks statistic train | Ks statistic test |
| 0.4219 | 0.3811 |

For training sample, the KS value is 0.4219, which is higher than the threshold of 0.4, indicating that discriminatory power of the model is good. For testing sample, the KS value is 0.3811, which is lower than the threshold of 0.4, indicating that discriminatory power of the model is not sufficient.

5.2.2 Variance Inflation Factor

This test is to assess the multicollinearity among the risk drivers. The variance inflation factor (VIF) quantifies the severity of multicollinearity in an regression analysis caused by correlation between multiple ‘independent variables’ in a model (i.e., the risk drivers). It is a measure of how much the variance of an estimated regression coefficient is inflated because of multicollinearity in the model. High levels of multicollinearity results in unstable parameter estimates and renders the model statistically invalid.

The test returns VIF value for each variable listed in the table below.

|  |  |
| --- | --- |
| Variable | Vif |
| priorfydscr\_transformed | 1.3352 |
| mrfytdocc\_transformed | 1.6708 |
| debt\_yield\_p1\_transformed | 1.364 |
| priorfyocc\_transformed | 1.6813 |
| OLTV\_transformed | 1.2259 |
| Division\_transformed | 1.0953 |
| interestonly\_transformed | 1.045 |

As shown in the above table, all variables have a VIF lower than 4, indicating the model does not suffer from multicollinearity.

5.2.3 Receiver Operating Characteristic curve

Receiver Operating Characteristic (ROC) is a curve in a unit square which is used to assess the accuracy of the diagnostic (e.g. rating) system. The area under curve (AUC) is the area under the ROC curve. Higher AUC means higher discriminatory power of model.

|  |  |
| --- | --- |
| Area under curve | Gini coefficient |
| 0.7743 | 0.5487 |

As shown in the above table, the AUC value is higher than 0.7 and the Gini value of the model is higher than 0.4, indicating that discriminatory power of the model is good.

5.2.4 Normal Test

The aim of the test is to evaluate the adequacy of observed default rates comparing to predicted PD. The test is performed on the level of rating grade. It is applied under the assumption that the mean default rate does not vary too much over time and that default events in different years are independent. The normal test is motivated by the Central Limit Theorem and is based on a normal approximation of the distribution of the time-averaged default rates. With high enough (>30) number (n) of observations which are independent, Point in Time (PIT) observed default rate (ODR) is expected to be normally distributed. The algorithm of the test is to calculate time-weighted (or simple arithmetic) average of all PiT ODRs, calculate standard error of this statistic and derive confidence interval for it.

The test returns a table of p values and a brief interpretation of the test result for each rating grade.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rating class | Predicted pd upper boundary | Default rate | Normal test p value | Normal test result |
| 1 | 0.1% | 0.13% | 0.1063 | Acceptable |
| 2 | 0.26% | 0.22% | 0.1439 | Acceptable |
| 3 | 0.42% | 0.42% | 0.1088 | Acceptable |
| 4 | 0.56% | 0.53% | 0.1982 | Acceptable |
| 5 | 0.71% | 0.7% | 0.2359 | Acceptable |
| 6 | 0.93% | 0.91% | 0.2364 | Acceptable |
| 7 | 1.12% | 0.7% | 0.2665 | Acceptable |
| 8 | 1.41% | 1.1% | 0.4014 | Acceptable |
| 9 | 2.56% | 1.91% | 0.4329 | Acceptable |
| 10 | 15% | 5.37% | 0.0627 | Acceptable |

As shown in the above table, the p-value of Normal test for all segments is higher than 0.05, which means the null hypothesis cannot be rejected and there is evidence that the predicted PD does not deviate from the long run average of observed values. Therefore, the accuracy of the model is adequate.