

# Validation of Credit Rating Models - A Preliminary Look at Methodology and Literature Review

Ming-Yi Sun, Szu-Fang Wang  
JCIC Risk Research Team

## I. Introduction

In preparing for the formal implementation of the New Basel Capital Accord (Basel II) at the end of 2006<sup>1</sup>, our banking sector has been studying relevant provisions and response strategies. In the hope to promote and keep our banking supervision and risk management at the international level, the Bureau of Monetary Affairs under the Financial Supervisory Commission (formerly the Bureau of Monetary Affairs under the Ministry of Finance) in particular has set up a New Basel Capital Accord Joint Research Taskforce with Bankers Association to study relevant regulatory and implemental issues. The banking sector is paying particular attention to the internal-ratings based (IRB) approaches for credit risk provided in Basel II. Especially, model validation has been the focus among practitioners, which plays an important role in IRB qualification by supervisor. As an introductory effort, this

paper tackles the subject of credit rating model validation. In reference to current theories and practices on the subject, we examine the considerations for model validation and introduce currently adopted approaches. However, readers should keep in mind that this paper only discusses quantitative approaches. New theories and approaches for qualitative validation will be discussed at a later date as this field of study develops. If a bank realizes the whole picture about model validation, it will facilitate the work of IRB model construction and strategic planning for its business operation. More so, if the rating system is accepted by the regulatory authority, it will certainly boost the bank's stature and market competitiveness.

Below is an introduction to the minimum operational requirements for the validation of IRB model outputs suggested in the draft of Basel II, complemented with actual case study. Hopefully it will provide some value to

---

<sup>1</sup> The Internal-Ratings Based approaches for credit risk and the Advanced Measurement Approaches (AMA) for operational risk will be implemented at the end of 2007.

banks that intend to adopt the IRB approach.

## II. Validation Framework

The primary purpose of validation is to examine whether the internally constructed scoring model can fully explain the credit status of borrowers. As sampled data used to construct the model can mostly be explained by the model, it is necessary to see whether the model possesses sufficient explanatory power for different samples. Thus out-sample testing should be carried out to observe the tendency of over-learning, which will lower the predictability of model. In addition, improper sampling and omission of relevant information will lead to model bias. Thus external data should be employed to assess the validity of the rating model. Moreover, since the primary objective of the rating model is to make forecast, whether the model works normally under all circumstances, including significant changes of the macroeconomic environment, must also be validated. Below is an introduction to the framework of model validation.

### A. Backtesting

Backtesting entails the use of out-samples, including samples of different periods not used in model construction, samples of the same periods not used in model construction, as well as samples of different periods used in model construction, to examine the out-sample predictability of model.

### B. Benchmarking

Benchmarking is to compare the results of rating individual borrower or facility using

internal models with the outputs of external mechanisms, analyzing the origin of disparity and deciding whether such disparity is reasonable. The benchmark could be market information (spread), or the assessment of third parties (e.g. external rating agency or external models) or internal model (other rating systems). Differing from backtesting which stresses discerning the discrepancy between prediction and actual outcome, benchmarking stresses the dissimilarities between different predictors.

### C. Calibration

If the model only takes account of all-over efficacy, it would hide the part of poor predictability. Thus, it has to review the all-over and each departmental efficacy. Calibration will advance the predictability of model by inspecting the deviations in all situations and adjusting them. To compare and determine that the estimates of risk components for each grade of internal rating are within reasonable range, banks can use historical data, the assessment of external rating agencies and external model outputs or the outputs of other internal rating systems. The comparison basis can be measures of risk components, expected losses or unexpected losses. The comparison can be made over one grade, multiple grades or the all asset portfolio to observe if there is any material discrepancy. Differing from benchmarking which observes whether the rating results are consistent, calibration targets risk weights in the same grade to see if they are consistent.

#### D. Stress test

Stress test entails simulating adverse economic conditions or expected events by means of trial calculation or scenario analysis to observe possible resulting changes and losses.

### III. Validation Methods

The validity of a rating model is judged from three dimensions: 1. discriminative power: the accuracy of the model in differentiating non-defaulters and defaulters; 2. homogeneous: does the model provide enough rating grades to classify borrowers with different credit characteristics, while the credit characteristics of all borrowers in the same grade are homogenous; 3. stability: a good rating model must take into account the influence of external economic factors that the model outputs reflect the credit status of individual borrowers and represent long-term trends without being affected by short-term volatility. The validation methods for different dimensions of a rating model are discussed below.

#### A. Analysis of discriminative power

To validate whether a credit rating model has adequate discriminative power and to examine whether its error is within a reasonably acceptable range, the following methods (but not limited to those methods) are recommended:

##### 1. Kolmogorov-Smirnov Test (K-S test)

The credit rating model has to discriminate the difference between non-defaulters and defaulters. Therefore, adoption of K-S test would validate whether the distribution of

rating score of non-defaulters differ with the distribution of rating score of defaulters and understand whether the credit rating model discriminate the difference between non-defaulters and defaulters.

The theoretical basis of K-S test is discussed below:

When the difference in the cumulative relative frequency distribution between two samples of data is very small and such difference is random, the population distribution of the two samples should be consistent; conversely, when the distribution of two populations is not consistent, the difference in the cumulative relative frequency distribution of the two samples will be significant, as can be seen in Figure 1.

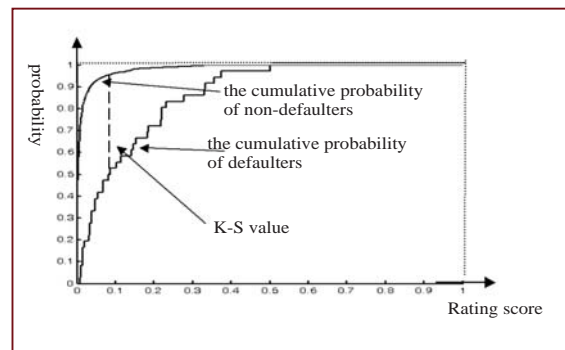


Figure 1: The cumulative probability distribution of non-defaulters and defaulters

- (1) Calculate the cumulative probability of non-defaulter and defaulters in each stage of rating score.
- (2) Calculate the difference in the cumulative probability between two groups in each stage.

(3) Find the maximum difference in the cumulative probability (K-S value).

After a credit rating model has been constructed, we can use K-S test to examine whether the ratings of non-defaulters and defaulters are uniformly distributed; the greater the K-S value, the more it proves that the rating scores of non-defaulters and defaulters are not uniformly distributed, the better the scoring model is in discriminating the difference between non-defaulters and defaulters.

## 2. Gini Coefficient

The rating of non-defaulters and defaulters would be different. Gini coefficient is essentially a variance used to quantify the difference between data points. In Figure 2, when the model would not discriminate the difference of non-defaulters and defaulters. Thus non-defaulters and defaulters will have the same distribution,

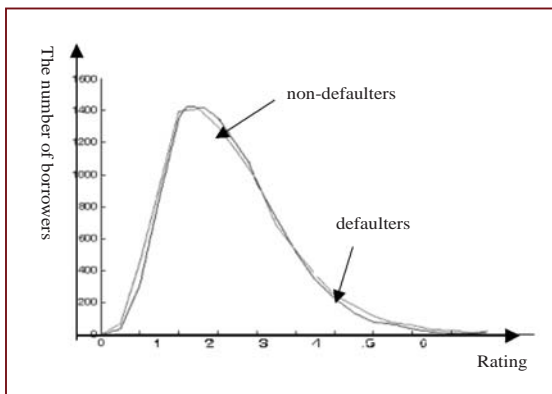


Figure2: The rating of non-defaulters and defaulters

In Figure 3, the x axis is the ratio of the borrowers to all borrowers below K; the y axis is the ratio of defaulters to all

defaulters below K. Lorenz curve is the segment of each point under each rating. The Lorenz curve will overlap with the 45 degree line in the chart as the model would not discriminate the difference of non-defaulters and defaulters. Thus the area between the Lorenz curve and 45 degree line is the measure of difference in the distribution of non-defaulters and defaulters. Gini coefficient is the quotient of this area divided by the entire area below 45 degree line as expressed by:  $gini = \frac{A}{A+B}$   $= \frac{A}{0.5} = 2A$ . That is, Gini coefficient is twice the area enclosed by the Lorenz curve and 45 degree line.

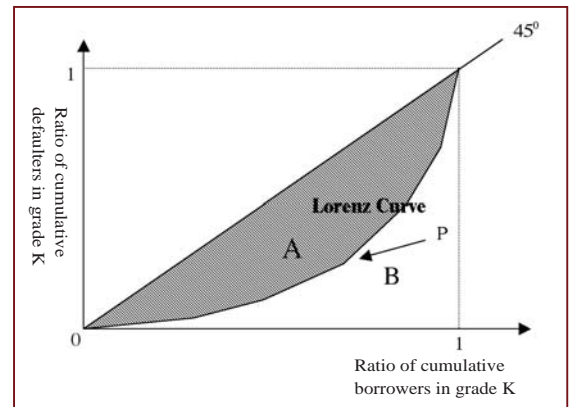


Figure3: Lorenz curve

Gini coefficient can be expressed by a mathematical equation:

$$G = \left| 1 - \sum_{K=0}^{K=n-1} (X_{K+1} - X_K)(Y_{K+1} + Y_K) \right| \quad (1)$$

$G$  : Gini coefficient

$X$  : Cumulative percentage of borrowers

$Y$  : Cumulative percentage of defaulters

$K$  : Rating

Gini coefficient ranges between 0 and 1; when it is equal to 1, it means the model outputs result in unequal distribution of non-defaulters and defaulters, thereby is fully able to differentiate defaulters and defaulters; when it is equal to 0, the rating model cannot create unequal distribution of the two groups, hence totally unable to differentiate between defaulters and defaulters.

### 3. Receiver Operating Characteristic (ROC)

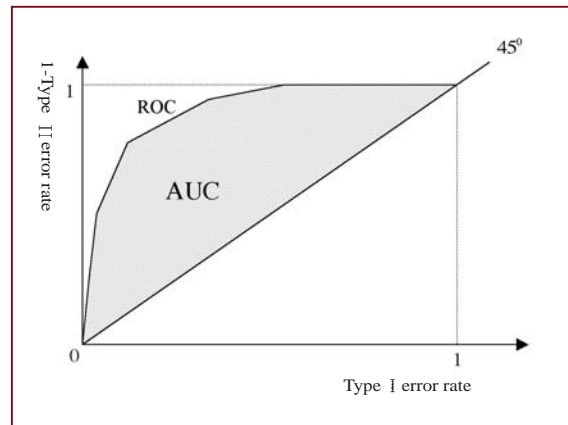
Let  $C$  is the cut point. When a bank classifies non-defaulters and potential defaulters based on the rating results, it is bound to incur Type I error (creditworthy borrower is classified as a potential defaulter) and Type II error (defaulting borrower is classified as a non-defaulter), as derived in Table 1.

**Table1: The decision under the critical value**

Cut point $C$		Defaulters	Non-defaulters
Rating score	Above critical value	Correct prediction	Type I error
	Below critical value	Type II error	Correct prediction

ROC curve is plotted based on Type I error rate (false alarm rate) and 1 minus Type II error rate (hit rate) under all possible decisions made by the decision-maker as shown in Figure 3:

The more the ROC curve bends towards (0, 1), the better the rating model is able to distinguish non-defaulters and defaulters. In other words, the bigger area under the ROC



**Figure4: ROC curve**

curve (area under curve or AUC), the more accurate the rating model. AUC can be interpreted as the average ability of the rating model to accurately classify non-defaulters and defaulters. When AUC is 0.5, it means non-defaulters and defaulters are randomly classified; when AUC is 1, it means the scoring model accurately classifies non-defaulters and defaulters. Thus in reality, the AUC ranges between 0.5 and 1 (AUC under 0.5 has no meaning).

We can use an unbiased estimator  $\hat{U}$  to denote AUC:

$$\hat{U} = \frac{1}{N_D \times N_{ND}} \sum_{(D,ND)} u_{D,ND} \quad (2)$$

$$u_{D,ND} = \begin{cases} 1 & \text{if } S_D < S_{ND} \\ \frac{1}{2} & \text{if } S_D = S_{ND} \\ 0 & \text{if } S_D > S_{ND} \end{cases}$$

where

- $N_D$  is the total number of defaulters;
- $N_{ND}$  is total number of non-defaulters
- $S_D$  is the rating score of defaulters;
- $S_{ND}$  is the rating score of non-defaulters.

#### 4. Cumulative Accuracy Profile (CAP)

A credit rating model would assign the lowest credit scores to the default. Consider a rating model which assigns to each debtor a scores  $s$  out of  $k$  possible values  $\{S_1, \dots, S_k\}$  with  $S_1 < \dots < S_k$ . A high rating score indicates a low default probability. By the scoring results, the cumulative percentage of all defaulters should be equal to the cumulative percentage of all borrowers multiplied by default rate. But in reality, such is oftentimes not the case. The discrepancy comes from the model error in interpreting non-defaulting and defaulting borrowers. The Cumulative Accuracy Profile curve is defined as the graph of all points  $(CD_T^i, CD_D^i)_{i=0, \dots, k}$  where the points are connected by straight lines as shown in Figure 4.<sup>2</sup>

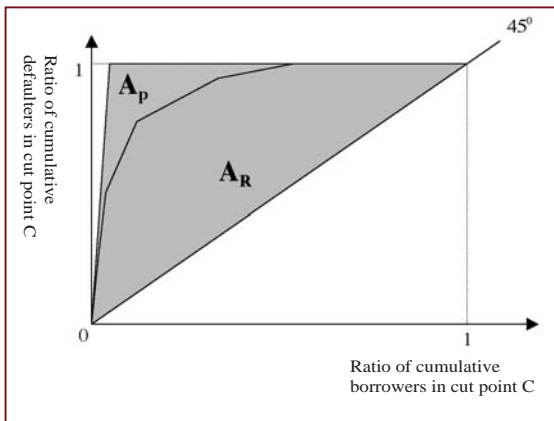


Figure5: CAP curve

In the best situation, the CAP curve would be a straight line with gradient of (1/default rate) and staying at 1. Conversely, the CAP curve of a model without any discriminative power would be a straight 45° line. In reality, the CAP curve of a rating model would run between the two. The quantification of CAP curve is termed “accuracy rate”, which is defined as the ratio of  $A_R$  (area between the CAP curve of scoring model and 45° line) and  $A_P$  (area between 45° line and the CAP curve of perfect model):

$$AR = \frac{A_R}{A_P} \quad (3)$$

AR lies between 1 and 0;<sup>3</sup> the closer it is to 1, the more accurate is the scoring model; conversely, the closer it is to 0, the less accurate is the model.

Generally we can use an unbiased estimator  $\hat{V}$  to denote AR:

$$\hat{V} = \frac{1}{N_D \times N_{ND}} \sum_{(D, ND)} v_{D, ND} \quad (4)$$

$$v_{D, ND} = \begin{cases} 1 & \text{if } S_D < S_{ND} \\ 0 & \text{if } S_D = S_{ND} \\ -1 & \text{if } S_D > S_{ND} \end{cases}$$

where  $N_D$  is total number of defaulters;  $N_{ND}$  is total number of non-defaulters;  $S_D$  is the credit score of defaulters;  $S_{ND}$  is the credit score of non-defaulters.

<sup>2</sup> We define the cumulative probabilities as  $CD_T^i$  and  $CD_D^i$  that denote the distribution function of the  $i$  score value of the total sample of debtors and defaulters.

<sup>3</sup> When each data points can be further divided, Gini coefficient is equal to accuracy rate (AR).

**B. Homogenous**

Assuming the credit risk model would discriminate correctly the difference of non-defaulters and defaulters, it should discriminate the borrowers have different creditworthy. The great grade discrimination would put the borrowers have the same creditworthy in the same grade. In other words, the difference of default factor of borrowers in the same grade is similar. Thus, the great grade discrimination would make difference be the smallest. Assuming K borrowers in the same rating grade have the same Probability of Default (PD), all defaults within the grade are binomially distribution, that the defaults per rating grade are statistic independent. In other words, all borrowers within the same rating grade must be homogenous. If not, it means default events are not independent of each other and the estimation of default rate for this rating grade is not accurate. Thus it is necessary to use binomial test to verify whether all borrowers within the same grade are statistically independent, the following methods (but not limited to those methods) are recommended:

**1. Binomial test**

When the number of default events ( $D_K$ ) in a rating grade containing K borrowers exceeds a critical value  $d_{k,a}$ , we can reject the hypothesis that the actual PD will be smaller than or equal to the estimated PD at a confidence level  $\alpha$ ; in other words, no

sufficient evidence shows that PD is underestimated. The  $d_{k,a}$  is calculated as follows:

$$d_{k,\alpha} = \min \left\{ d : \sum_{i=d}^K \binom{K}{i} PD^i (1 - PD)^{K-i} \leq \alpha \right\} \quad (5)$$

In light that binomial test ignores the effects of economic fluctuation and asset correlation, it generally underestimates.  $d_{k,a}$  In the calibration of PD, binomial test provides a conservative indicator.

**2. Granularity adjustment**

Given that binomial test omits the influence of asset correlation, resulting in underestimation of critical value  $d_{k,a}$ , we attempt to add in this factor to relax the constraint of  $d_{k,a}$ .

According to the hypothesis of one factor model of Gordy (2002): defaults are influenced by systematic and idiosyncratic factors. The idiosyncratic factors are mutual independent. All borrowers are influenced by the same systematic factors. The influences of systematic factors are the correlation ( $\rho$ ) of assets. Therefore, critical value  $d_{k,a}$  may be simplified as:<sup>4</sup>

$$d_{k,\alpha} \approx k \times q(\alpha, R) + \frac{1}{2} \left\{ 2 \times q(\alpha, R) - 1 + \frac{q(\alpha, R)(1 - q(\alpha, R))}{\phi \left( \frac{\sqrt{\rho} q(\alpha, R) - t}{\sqrt{1 - \rho}} \right)} \right\} \quad (6)$$

<sup>4</sup> As derived in appendix A1

$$\left[ \frac{\sqrt{\rho}q(1-\alpha, X)-t}{\sqrt{1-\rho}} - \sqrt{\frac{1-\rho}{\rho}}(q(1-\alpha, X)) \right] + 1$$

### 3. Moment matching

Granularity Adjustment assume probability of default following Normal-distribution, thus critical value  $d_{k,\alpha}$  may be simplified as (6). According to the hypothesis of Overbeck and Wanger(2000): using Moment Matching, probability of default follows Beta-distribution. Therefore, critical value  $d_{k,\alpha}$  may be simplified as:<sup>5</sup>

$$d_{k,\alpha} = k \times q(\alpha, Z) + 1 \quad (7)$$

Where Z is random variable and the probability density function of Beta-distribution is given by

$$\beta(a, b, x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \quad (8)$$

Then, we use the data of the Joint Credit Information Center (JCIC) to validate the homogeneous in the same grade. When the true defaulters exceed the critical value, the results show that the borrowers reject the hypothesis of homogeneous. In Table 2, the homogeneous of out-sample is validated. Percent is the ratio of borrowers in each grade to total sample. PD is the estimated default rate in each grade. No. of borrower is the number of borrowers in each grade. Binomial is the tolerance of defaulters in each grade by Binomial Test. Granularity is the tolerance of defaulters in each grade by Granularity Adjustment. Moment is the tolerance of defaulters at each grade by Moment Matching. Default is true defaulters in each grade. DR is the true default rate in each grade.

Table 2: Validation of Grade Homogeneity

Rating	Percent	PD	No. of borrower	Binomial	Granularity	Moment	Default	DR
1	1.59%	0.00%	1651	1	1	1	1	0.06%
2	2.99%	0.93%	3103	47	423	329	7	0.23%
3	3.46%	2.26%	3601	110	724	609	17	0.55%
4	10.46%	2.54%	10869	328	2286	1938	74	0.68%
5	14.78%	3.03%	15362	532	3479	2992	144	0.94%
6	23.32%	4.80%	24235	1267	6761	6025	388	1.60%
7	17.81%	8.25%	18513	1644	6865	6339	484	2.61%
8	16.39%	17.98%	17036	3219	9635	9278	840	4.93%
9	9.20%	40.20%	9566	3994	7749	7710	1155	12.07%

<sup>5</sup> As derived in appendix A2



**4. CIER (condition information entropy ratio)**

Assuming all rating grades output by a credit rating system correspond to a PD value and assuming this system can output  $n$  rating grades  $R = \{ R_1, \dots, R_n \}$ , the conditional information entropy can be expressed as follows:

$$H_1(R) = - \sum_k P(R_k) (P(D|R_k) \times \log P(D|R_k) + P(N|R_k) \times \log(N|R_k)) \quad (9)$$

where

$P(D|R_k)$  is the probability of default among borrowers in rating grade  $R_k$ .

$P(N|R_k)$  is the probability of non-default among borrowers in rating grade  $R_k$ .

A good credit rating system can differentiate borrowers having different credit quality and assign corresponding grades. In practice, Conditional Information Entropy Ratio (CIER) as defined below is used to evaluate the quality of grades assigned by a rating system to borrowers of different credit conditions.

$$CIER = \frac{H_0 - H_1(R)}{H_0} \quad (10)$$

Assume default events obey binomial distribution, then

$$H_0 = H_1(p) = -p \log(p) - (1-p) \log(1-p) \quad (11)$$

$P$  denotes the PD of all borrowers.

A perfect credit rating system is able to classify potential defaulters in one grade

and non-defaulters in other grades. In such event, CIER is equal to 1, for the system totally predicts the uncertainty of default information. For a rating system that does not have any discriminative power, the distribution of potential defaulters in each grade will be identical to that of the population, meaning the system cannot provide additional information and its CIER is 0.

In Table 3, the results would be validated by CIER. The data based on the JCIC and Sobehart, Keenan & Stein (2001).

**Table 3: Validation of Grade Distribution**

ROA	0.06
Reduced Z-score	0.09
Z-score	0.06
Hazard Model	0.11
Merton Model Variant	0.14
Mood's model	0.19
JCIC	0.1

**C. Stability analysis**

Stability analysis pertains to observing whether the model results show the short-run and long-run trends of drastic change, and further, analyzing the impact of short-term economic fluctuation on the basis of long-term rating, variation of grades resulting from change of rating method, whether grade changes comply with basic model assumptions or a manifestation of model deficiency, and thereby analyzing whether the change of transition matrix is within reasonable range.

## 1. The establishment of transition matrices

Analyzing the stability of model have to know the change of each rating, and future, understanding whether the range of rating change are reasonable. These results would be analyzed whether the model results show the short-run and long-run trends of drastic change. Thus, the transition probability of each grade will influence the analysis of stability of model. the following methods (but not limited to those methods) are recommended:

### a. Cohort approach

Let  $P_{i,j}(\Delta_t)$  be the probability of migrating from grade  $i$  to  $j$  over horizon  $\Delta_t$ . E.g. for  $\Delta_t=1$  year, there are  $n_i$  borrowers in rating category  $i$  at the beginning of the year, and  $n_{i,j}$  migrated to grade  $j$  by year-end. Then, an estimate of the transition probability.

$P_{i,j}(\Delta_t = 1) = \frac{n_{i,j}}{n_i}$ . In general, using this approach, the borrowers whose rating were

withdrawn or migrated to not rated status are removed form the sample.

### b. Duration approach

For a time homogeneous Markov chain, the transition probability matrix is a function of the distance between dates. The estimates of the elements of the intensity matrix  $\gamma$  is given by

$$\hat{\gamma}_{i,j} = \frac{n_{i,j}(T)}{\int_0^T Y_i(S) ds} \quad (12)$$

where  $Y_i(s)$  is the number of borrowers with rating  $i$  at time  $S$ , and  $n_{i,j}(T)$  is the total number of transitions over the period form  $i$  to  $j$ . Moreover, the Markov transition probability matrix with time homogeneous would be obtained by exponential of intensity matrix.

For example, transition matrix is established in Table 4.<sup>6</sup>

Table 4: Transition matrix

End Start	AAA	AA	A	BBB	BB	B	CCC	default
AAA	91.93%	7.46%	0.48%	0.08%	0.04%	0.00%	0.00%	0.00%
AA	0.64%	91.81%	6.76%	0.60%	0.06%	0.12%	0.03%	0.00%
A	0.07%	2.27%	91.38%	5.12%	0.56%	0.25%	0.01%	0.04%
BBB	0.04%	0.27%	5.56%	87.87%	4.83%	1.02%	0.17%	0.24%
BB	0.04%	0.10%	0.61%	7.75%	81.48%	7.90%	1.11%	1.01%
B	0.00%	0.10%	0.28%	0.46%	6.95%	82.80%	3.96%	5.45%
CCC	0.19%	0.00%	0.37%	0.75%	2.43%	12.13%	60.44%	23.69%

<sup>6</sup> S&P's average one year transition rates, adjusted for the "not rated" category, date based on Cynthia & Ron (2000).

**2. Analyzing the reasonableness of transition matrix change (grade maintenance rate, rate of upgrade/downgrade, etc.)**

According to the approaches of establishment of transition matrix, we obtain a long-term credit transition matrix. Then, each grade change of transition matrix, especially drastic change pertains to analyzing whether the probability of grade change is decreasing with more drastic change and reasonable.

**3. Analyzing the reasonableness of grade reinstatement (grade reinstatement rate, etc.)**

The prediction of three year credit risk model would be reviewed to comparing whether the results is change in the same way. The prediction in different way is called reverse rating. Let  $n_{ij,jk}$  be the number of borrowers with migrating from grade  $i$  to  $j$  and to  $k$  ( $i > j$  and  $j < k$  or  $i < j$  and  $j > k$ ) and  $n_i$  be the total borrowers that initial grade is  $i$ . Reverse rating is given by

$$R_i = \frac{n_{ij,jk}}{n_i} \quad i = 1, \dots, N \tag{13}$$

The range of grade change would be modified by different situation. However, as the rage of grade change larger, the reverse rating would be smaller.

**4. Does cumulative PD vary monotonically with time and grade?**

The cumulative default probability each year would be reviewed to analyzing

whether the cumulative default probability is increasing with time and decreasing with downgrade.

**5. Singular value decomposition (SVD) of mobility**

Stability analysis must take account of the time homogenous of transition matrix to analyzing whether the model results the impact of short-term economic fluctuation on the basis of long-term rating. First, the identity matrix  $I$  is defined as homogenous matrix. This matrix means the rating of said borrower does not change with time. An actual matrix  $P$  with a distance from the homogenous matrix is defined as a mobile matrix  $\tilde{P}$ :

$$\tilde{P} = P - I \tag{14}$$

Y.Jafry and T.Schuermann (2004) propose a metric defined as the average singular value of a mobile matrix  $M_{svd}$ , described as follows, with the notion of mobility matrix and adoption of singular value decomposition.<sup>7</sup>

$$M_{svd}(P) = \frac{\sum_{i=1}^n \sqrt{\lambda_i(\tilde{P}'\tilde{P})}}{N} \tag{15}$$

where  $\lambda_i$  denotes the  $i^{th}$  eigenvalue.

For example, we obtain SVD value is 0.1563 by using the Table 4 that is S&P's average one year transition rates, adjusted for the “not rated” category.

<sup>7</sup> As derived in appendix A3

#### IV. Conclusion

Credit rating model validation covers extensive dimensions, each one of them is hard pressed to take into account all situations. Hence a mix of validation approaches should be more appropriate. Also as statistical figures cannot set a so-called reasonable range, benchmarking using external information and even different models but identical samples should be employed to examine the soundness of internal models. In model construction, sometimes it is difficult to match all measures at the same time (e.g. the model results have optimum discriminative power, stability and grade distribution at the same time). Thus banks should identify the primary objective of constructing an internal model and prioritize all the dimensions to make sure their internal model achieves the intended results.

#### Appendix A1

Using the one factor model of Gordy (2002), let  $\sqrt{\rho}X + \sqrt{1-\rho}\zeta_i \leq t$  depict a default event, then the total number of default is  $D_K$

$$D_K = \sum_{i=1}^K 1_{\{\sqrt{\rho}X + \sqrt{1-\rho}\zeta_i \leq t\}}$$

Define the defaults per rating grade are binomially distribution

$$E[D_k] = kp$$

Assume Probability of Default (PD) observes normal distribution, we can derive

$$t = \Phi^{-1}(p)$$

Let critical value  $d_{k,a}$  be

$$d_{k,a} = \min \{d: P[D_k \geq d] \leq 1 - \alpha\}$$

then  $q(\alpha, D_k)$ , the  $\alpha$  quantile of  $D_K$  is

$$q(\alpha, D_k) = \min \{x: P[D_k \leq x] \geq \alpha\}$$

Therefore, critical value  $d_{k,a}$  may be simplified as

$$d_{k,a} = q(\alpha, D_k) + 1$$

Because  $R_K = \frac{Dk}{K}$ , the quantile of  $R_K$  and  $D_K$  are related by

$$q(\alpha, R_K) = \frac{q(\alpha, D_K)}{K}$$

Use Taylor expansion to carry out second-order expansion of  $\alpha$  quantile of  $R_K$  at  $R$

$$\begin{aligned} q(\alpha, R_K) &= q(\alpha, R + h(Rn - R)) \Big|_{h=1} \\ &\approx q(\alpha, R) + \frac{\partial}{\partial h} q(\alpha, R + h(Rn - R)) \Big|_{h=0} \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial h^2} q(\alpha, R + h(Rn - R)) \Big|_{h=0} \end{aligned}$$

when  $K$  approximates infinity,

$$\lim_{K \rightarrow \infty} R_K = R = \Phi \left( \frac{t - \sqrt{\rho}X}{\sqrt{1-\rho}} \right)$$

The  $\alpha$  quantile of  $R$ ,  $q(\alpha, R)$  is

$$q(\alpha, R) = \Phi \left( \frac{\sqrt{\rho}\Phi^{-1}(\alpha) + t}{\sqrt{1-\rho}} \right)$$

After obtaining  $q(\alpha, R)$ , we can simplify the  $\alpha$  quantile of  $D_K$ ,  $q(\alpha, R)$  as

$$q(\alpha, D_k) \approx k \times q(\alpha, R) + \frac{1}{2} \left\{ 2 \times q(\alpha, R) - 1 + \frac{q(\alpha, R)(1 - q(\alpha, R))}{\phi\left(\frac{\sqrt{\rho}q(\alpha, R) - t}{\sqrt{1 - \rho}}\right)} \left[ \frac{\sqrt{\rho}q(1 - \alpha, X) - t}{\sqrt{1 - \rho}} - \sqrt{\frac{1 - \rho}{\rho}}(q(1 - \alpha, X)) \right] \right\}$$

Subsequently, critical value  $d_{k,\alpha}$  that takes into account asset correlation can be derived.

### Appendix A 2

Assuming PD follows beta-distribution, the  $\alpha$  quantile of  $D_k$ ,  $q(\alpha, D_k)$  may be rewritten as

$$q(\alpha, D_k) \approx k \times q(\alpha, z)$$

The density of a  $B(a,b)$ -distributed random variable  $Z$  is defined by

$$\beta(a, b, x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$$

Its parameters  $a$  and  $b$  are respectively

$$a = \frac{E[R_k]}{\text{var}[R_k]} (E[R_k] (1 - E[R_k]) - \text{var}[R_k])$$

$$b = \frac{1 - E[R_k]}{\text{var}[R_k]} (E[R_k] (1 - E[R_k]) - \text{var}[R_k])$$

where

$$E[R_k] = p$$

$$\text{var}[R_k] = \frac{K-1}{K} \Phi_2(t, t, \rho) + \frac{p}{k} - p^2$$

Using Taylor expansion, we can derive

$$\Phi_2(t, t, \rho) \approx \Phi(t)^2 + \frac{e^{-t^2}}{2\pi} \left( \rho + \frac{1}{2} \rho^2 t^2 \right)$$

The expectation and variance respectively of  $Z$  are given by

$$E[Z] = \frac{a}{a+b}$$

$$\text{var}[Z] = \frac{ab}{(a+b)^2(a+b+1)}$$

Then we can obtain the  $\alpha$  quantile of  $Z$  ( $q(\alpha, R)$ ), and critical value  $d_{k,\alpha}$  is

$$d_{k,\alpha} = k \times q(\alpha, Z) + 1$$

### Appendix A 3

We can use this approach to find continuously mobile matrix according to the steps below:

Assuming  $A$  is a  $m$ -by- $n$  multiple matrix, and in matrix  $A^*A$ ,  $A^*$  is a transition of  $A$ , take the complex conjugate of each element, i.e.  $A^*_{ij} = \overline{A_{ji}}$ . If  $A^*=A$ ,  $A$  is termed a “hermitian matrix” with eigenvalue being a real number. Apparently  $A^*A$  is also hermitian. So its eigenvalue is a real number and non-negative.

The singular value decomposition of a  $m$ -

by-n matrix  $A$  is to break  $A$  into  $A = USV^*$ , where  $U$  is  $m$ -by- $m$ ,  $V$  is a  $n$ -by- $n$  unitary matrix,  $S$  is a  $m$ -by- $n$  non-standard "diagonal matrix",  $S_{ij} = 0, i \neq j$ ,  $S_{ii} = \sigma_i$ ,  $\sigma_i$  is a nonnegative real number. We call  $\sigma_i$  the singular value of  $A$ , which is the square root of eigenvalue of  $A^*A$  matrix.

Furthermore,  $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_s)$ ,  $s = \min(m, n)$ , and  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_s \geq 0$ . The feature vectors of  $A^*A$  constitute  $V$ , the feature vectors of  $AA^*$  constitute  $U$ , then:

$AV = US$ , i.e.  $Av_i = \sigma_i u_i, 1 \leq i \leq \min(m, n)$ ,  $u_i$  and  $v_i$  are respectively the column vector of  $U$  and  $V$ , and  $A^*A = VS^*U^*USV^* = V(S^*S)V^*$ , or  $S^*S = V^*(A^*A)V$ ,  $S^*S = D$  is a  $n$ -by- $n$  diagonal matrix, where main diagonal elements are  $\sigma_i^2, i=1, 2, \dots, s$ , while the rest are zero. Similarly  $AA^* = USV^*VS^*U^* = U(SS^*)U^*$ , or  $SS^* = U^*(A^*A)U$ ,  $SS^* = D'$  is a  $m$ -by- $m$  diagonal matrix, where the main diagonal elements are  $\sigma_i^2, i=1, 2, \dots, s$ , while the rest are zero.

A 3-D credit migration matrix is cited as an example to describe the use of SVD:

Assuming a credit migration matrix as follows:

$$P = \begin{bmatrix} 1-p_1 & p_1 & 0 \\ 0 & 1-p_2 & p_2 \\ 0 & p_3 & 1-p_3 \end{bmatrix}$$

then  $\tilde{P}$  and  $\tilde{P}^*$  are respectively:

$$\tilde{P} = \begin{bmatrix} -p_1 & p_1 & 0 \\ 0 & -p_2 & p_2 \\ 0 & p_3 & -p_3 \end{bmatrix}$$

and

$$\tilde{P}^* \tilde{P} = \begin{bmatrix} p_1^2 & -p_1^2 & 0 \\ -p_1^2 & p_1^2 + p_2^2 + p_3^2 & -(p_2^2 + p_3^2) \\ 0 & -(p_2^2 + p_3^2) & (p_2^2 + p_3^2) \end{bmatrix}$$

the eigenvalue of  $\tilde{P}^* \tilde{P}$  is:

$$\begin{bmatrix} p_1^2 + p_2^2 + p_3^2 + \sqrt{p_1^4 + p_2^4 + p_3^4 - p_1^2(p_2^2 + p_3^2) + 2p_2^2 p_3^2} \\ p_1^2 + p_2^2 + p_3^2 - \sqrt{p_1^4 + p_2^4 + p_3^4 - p_1^2(p_2^2 + p_3^2) + 2p_2^2 p_3^2} \\ 0 \end{bmatrix}$$

So the norm of  $\tilde{P}$ , i.e. the square root of the largest eigenvalue of  $\tilde{P}^* \tilde{P}$  is:

$$\|\tilde{P}\| = \sqrt{p_1^2 + p_2^2 + p_3^2 + \sqrt{p_1^4 + p_2^4 + p_3^4 - p_1^2(p_2^2 + p_3^2) + 2p_2^2 p_3^2}}$$

$$\|\tilde{P}x\|_{\max} = \|\tilde{P}\| \|x\|_{\max}$$

$$x'_{\max} = \begin{bmatrix} p_1^2 - p_2^2 - p_3^2 + \sqrt{p_1^4 + p_2^4 + p_3^4 - p_1^2(p_2^2 + p_3^2) + 2p_2^2 p_3^2} \\ p_2^2 + p_3^2 - \sqrt{p_1^4 + p_2^4 + p_3^4 - p_1^2(p_2^2 + p_3^2) + 2p_2^2 p_3^2} \\ 1 \end{bmatrix}$$

If  $p_1 = p_2 = p_3 = 0.1$ , then  $x'_{\max} = (-0.01 \ 0.02 \ 1)$

The SVD of this matrix is equal to

$$M_{svd} = \frac{\sqrt{2}}{3} \sqrt{p_1^2 + p_2^2 + p_3^2 + p_1} \sqrt{3(p_2^2 + p_3^2)}$$

## Reference

Altman and Rijken (2004) "How rating agencies achieve rating stability."

"Approaches to the Validation of Internal Rating System", DEUTSCHE BUNDESBANK Monthly Report September 2003

[http://www.bundesbank.de/download/volkswirtschaft/mba/2003/200309\\_en\\_rating.pdf](http://www.bundesbank.de/download/volkswirtschaft/mba/2003/200309_en_rating.pdf)

- Cynthia McNulty and Ron Levin(2000) “Modeling Credit Migration” RISK,February 2000
- Engelmann B., Hayden E., Tasche D.(2003) “Measuring the Discriminative Power of Rating System” , Discussion Paper Series 2: Banking and Financial Supervision, No. 01/2004, Deutsche Bundesbank.
- Gunter Loeffler (2004) “An anatomy of rating through the cycle, Journal of Banking & Finance(28)” , p695-p702.
- Jafry & Schuermann (2004)” Measurement Estimation and Comparison of Credit Migration Matrics.”
- Sobehart J., Keenan S., Stein R (2001) “Benchmarking Quantitative Default Risk Model : A Validation Methodology.”  
<http://riskcalc.moodysrms.com/us/research/crm/53621.pdf>
- Tasche D.(2003) “A Traffic Light Approach to PD Validation” , Working paper, [http://www-m4.ma.tum.de/pers/tasche/traffic\\_lights.pdf](http://www-m4.ma.tum.de/pers/tasche/traffic_lights.pdf)