Technique Document for Large Corporate Scorecard

Vietnam Prosperity Joint-Stock Commercial Bank

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Author: Hendra Tan
Creation Date: 2014/04/16
Last Modification Date:
Version: 3.0
Status: Final draft

Document Control

Change Record

Date	Author	Version	Change Reference
2014/4/16	Vincent Ma	1.0	First Draft
2014/5/30	Vincent Ma	2.0	Second Draft
2014/6/30	Vincent Ma	3.0	Final Draft

The Change Record section of the document should identify the date of the initial release of the document as well as any additional versions. When possible the change record should indicate the primary focus for changes within each new version. The change reference should be used to indicate what section was changed as well as the content.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

TABLE OF CONTENTS

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

1. BACKGROUND

1.1. CURRENT BPF MODEL FOR SME/MSMES CUSTOMERS

The current scorecard model for VPBank used for BPF products has the structure as follows:

Ratio	Formula	Weight (abs. value)
Leverage	Total Liabilities / Total Equity	30%
Profitability	Profit after tax (2,1 or 0)	15%
Inventory Turnover	Inventory * 365 / Cost of Goods Sold	10%
Receivable Days	Accounts Receivable * 365 / Sales	10%
DSCR	(PAT + Depreciation + Interest) / 35% * Total Debt	25%
Working Capital	(Cash + Receivable + Inventory) / Current Liability	10%

Table 1: Current BPF Model for SME/MSMEs customers

The current BPF model is designed for SME and MSMEs whom are guaranteed by real estate by 100%. Moreover, the customer is eligible to apply only when customer currently has no loan or debt in group 3-5 under SBV loan categories.

The current BPF model covers measurement of the business performance in areas of ***Leverage, Profitability, Efficiency, Debt Service Coverage*** and ***Liquidity*** only. And for Profitability, Receivable Days and DSCR ratio, current BPF model required two year financial statement item to be able to calculate the ratio under specified formula.

1.2. PERFORMANCE OF CURRENT BPF MODEL

Model performance are measured by several statistics such as **Accuracy Ratio** (**AR**, also called **Gini Coefficient**, see Session 2.4 for definition), **Kolmogorov-Smirnov statistic (K-S),** etc. Here in the following session, we will only consider AR (Gini coefficient) as the key performance measurement statistic to evaluate the predictability of the credit scoring model.

The following tables show the performance of the BPF model for SME and MSMEs customers for reference. One-year financial statement data are used, the predictability is measured by AR:

Variables	AR	Original Weight	Optimized Weight

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	(Continues Case)	(abs. value)	(abs. value)
Leverage	-19.22%	30%	22.74%
Profitability	3.46%	15%	2.7%
Inventory Turnover	-12.73%	10%	10.51%
Receivable Days	-22.15%	10%	40.63%
DSCR	16.34%	25%	2.57%
Working Capital	13.23%	10%	20.86%

Please note that the result is based on development sample data, and the result is based on one-year financial statement data, which means Profitability, Receivable Days and DSCR ratio are calculated in one-year basis. For the reasons why we use one year, please refer to Session 2.1.2.

Table 2: BPF Model Performance for SME Segment (AR: 28.06%)

Variables	AR (Continues Case)	Original Weight (abs. value)	Optimized Weight (abs. value)
Leverage	-3.32%	30%	0.06%
Profitability	4.78%	15%	18.35%
Inventory Turnover	-22.74%	10%	33.79%
Receivable Days	-23.94%	10%	37.07%
DSCR	13.89%	25%	5.31%
Working Capital	7.20%	10%	5.42%

Table 3: BPF Model Performance for MSMEs Segment (AR: 23.68%)

We see that the current BPF model's predictability is at the unacceptable level since the general rule of thumb for expert model is that the AR should be above 30%. Nevertheless, the above results are for reference only, we will develop a new model for Large Corporations which we expect to give higher predictability than the referencing level of the current BPF model.

1.3. SCOPE OF NEW DEVELOPED MODEL

Although the current BPF model is applied to SME/MSMEs but not Large Corporations, we will base on the structure of this model, make our enhancement to develop the new model for Large Corporations with high enough predictability. To achieve this goal, our approach is to find some better candidates of financial ratio as well as extra information (e.g. CIC information) which perform well or have a better discriminatory power to replace the old variables or to add into the current model.

The financial ratios or additional information we choose must satisfy the following:

- The ratio is meaningful
- The ratio is commonly used as indicator of Good/Bad in business

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

- The ratio is used by any other bank across the world as one of the candidate in credit scoring model
- The ratio or additional information does not violate any policies or requirements

Moreover, we ask for VPBank's feedback on our list of financial ratios to make sure all the candidates are acceptable to be included in the model, if in case VPBbank disagreed or had doubt using some particular ratios, we would exclude those financial ratios from our model.

We try to make sure the new model cover the key aspects for examining the business's performance of a customer, and the score obtained from a model will reflect the customer's probability of default in the future.

2. MODEL DESIGN

2.1. DATA STRUCTURE

2.1.1. FINANCIAL/INCOME STATEMENT & CIC INFORMATION

The data provided by credit information center is on a yearly basis. The structure of the data looks like the following:

MACIC	THOIDIEM	CT_100	CT_110	CT_111	CT_112	CT_120	CT_121	CT_129	CT_130	CT_131
0101000000370000	20101231	9396000000	1939000000	1939000000	0	0	0	0	3393000000	3355000000
0101000001510000	20101231	31258000000	1257000000	1257000000	0	0	0	0	1256000000	1256000000
0101000001590000	20101231	17744000000	2245000000	2245000000	0	0	0	0	3807000000	3457000000

Table 4: Illustration of data structure

The data provided is at customer level, each customer is associated to a unique customer ID (e.g. MACIC) and the values of the financial/income statement items. For details of definition of different attributes, please find Appendix A for data dictionary of financial/income statement and Appendix B for data dictionary of CIC information.

2.2. MODEL STRUCTURE

The new Large Corporate scorecard model will consist of two parts. The first part is **Financial Ratios**, the second part is **CIC Information.** Financial Ratios part consists of financial ratios that are calculated based on the financial/income statement items. CIC information part includes the other CIC variables.

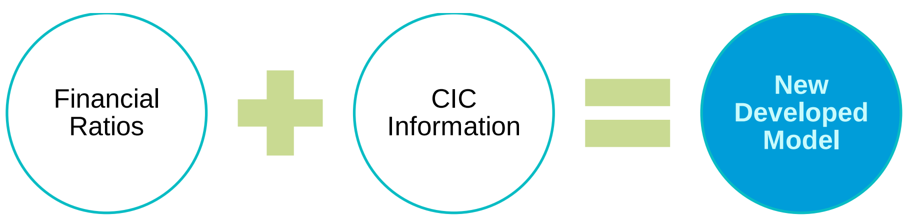


Figure 1: Illustration of model structure

For details about financial ratios and CIC information, please refer to Session 3.3.2.

2.2.1. ISSUE OF USING ONE-YEAR STATEMENT DATA OR TWO-YEAR

During model design phase, we have got an issue regarding using two-year financial/income statement item to calculate the financial ratios or just use one-year data. This argument occurs because for the current BPF model, there are some financial ratios which require two-year statement items to be calculated. In our perspective, even though more financial/income

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

statement data will allow us to calculate more feasible financial ratios, we propose using one-year statement data based on the following reasons:

I. Items reported on the statements are highly correlated for two consecutive financial years
II. Scarcity of Bad customers in our development sample

➢ **Items reported on the statements are highly correlated for two consecutive financial years**

It is common that financial statement or income statement data are highly correlated between consecutive financial years. To justify this claim, we perform **Spearman correlation analysis** to test whether the items in the financial statement and income statement are highly correlated between the year of application of the loan (Year N) and the previous year (Year N-1).

The result shows that more than **85%** of the items are highly correlated between year N and year N-1. For complete result please check the worksheet below:

Spearman Correlation.xlsx

Worksheet 1: Spearman correlation analysis

➢ **Scarcity of Bad customers in our development sample**

Another concern for our development sample size is whether the number of Bad customers is enough. The situation of our development sample for Large Corporations is as below:

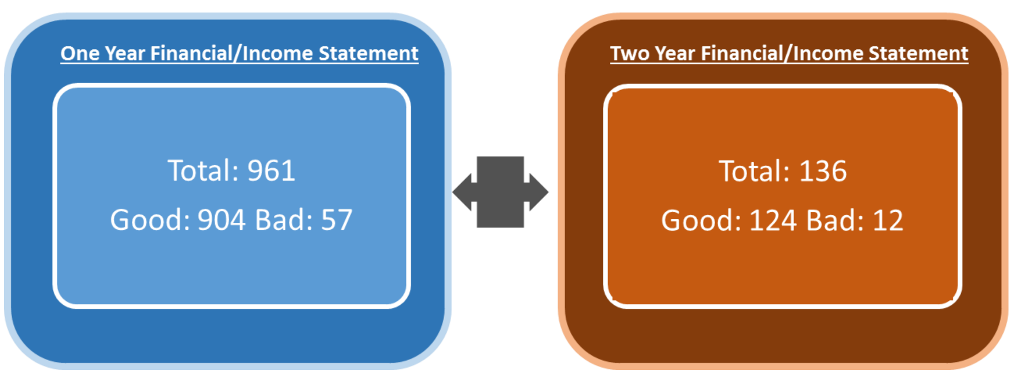


Figure 2: Comparison between using one-year financial/income statement data and two-year

If we use two-year financial/income statement, our development sample size will drop significantly from 961 to 136. The number of Bad customers drops from 57 to 12. The sample size and number of Bad customers for one-year statement data are already merely enough for developing a new scorecard, clearly the two-sample data will be unacceptable.

2.3. MODEL ASSUMPTION

All the data are extracted from CIC internal database. In case there are two different types of financial report (Tax or Internal), we assume that most of the reports summited to CIC are the internal reports from the banks. Therefore, when applying the scorecard, VPBank should also use internal reports rather than tax reports.

2.4. INTRODUCTION TO LOGISTIC REGRESSION & ACCURACY RATIO

There are many statistical methods that people apply in credit scoring models, such as linear discriminant analysis (LDA), k-nearest neighbor classification, classification and regression tree, neural networks, etc. But one of the most popular methods in practice is the logistic regression model. It is because of the intuitiveness of the model, clear interpretation of the coefficients, and the clarity and transparency of the mechanism how the risk factors function on the default stochastic. Put it plainly, the risk factors function on the default mechanism by adding linear effects to the log-odds, the coefficients reflect the direction and strength of the effects, and we can get very nice interpretation to the coefficients like odds-ratio, the interpretation of the model is well-understood by common statistical philosophy and terminologies, and there are rich and elegant theory, beautiful results and many insightful understandings can be established on such a model with clear and simple structure. These merits are not easily seen on other methods and it explains the popularity of logistic regression in industry and academia.

The logistic regression model has the following form:

$$logit(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_i \beta_i x_i$$

where p denotes PD, i.e. the probability of default, x_i are the (transformed) risk factors, β_0 is the intercept, β_i are the regression coefficients. The response/dependent variable is the binary outcome of default, i.e.Good=;0, Bad=1. The intercept and coefficients are measured by maximum likelihood estimates. After fitting the model (measuring the coefficients), we will be able to estimate the probability of default from the input of the risk factor values, in our scorecard model, these risk factors will be the customers' financial ratios and CIC information.

After the model is fitted, we need to evaluate how good the model is to discriminate between Good and Bad customers by the predicted PDs (or by the scores which are mapped from the regression coefficients in the model, and when the predicted PD is higher, the score should be lower), this ability of the model is called **Discriminatory Power**. A common statistical measure of the model's discriminatory power is **Accuracy Ratio (AR, or Gini Coefficient)**, which is defined from the **Cumulative Accuracy Profile (CAP)** curve as below:

Copyright © 2014 EntroFine. All Rights Reserved.

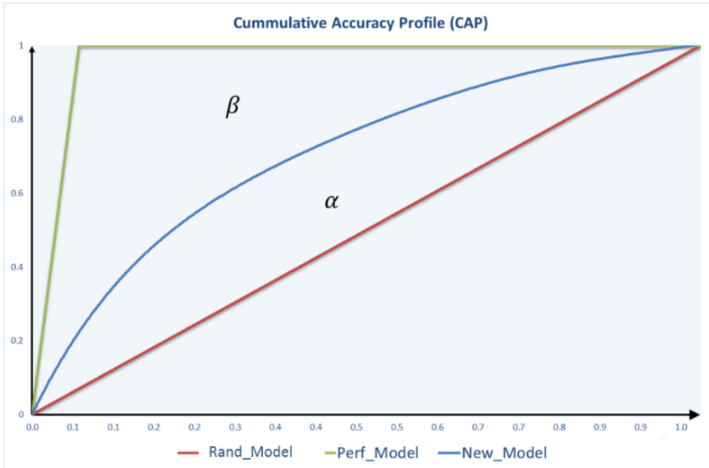


Figure 3: Cumulative accuracy profile (CAP) curve

Consider an arbitrary rating model that produces a rating score, where a high score is an indicator of a low default probability. The CAP curve is constructed by first sorting all debtors' scores in ascending order (lowest to highest), then at each score level, calculate the cumulative percentages in the defaulted customers and in the whole sample respectively, then connecting all the points of the pairs of two different cumulative percentages formed by all score levels will give the CAP curve. The rationale of CAP curve is that for a good scorecard, the lower scores are associated with higher probabilities of default. In the above figure the green line (Perf_Model) is the perfect model which accumulates all the bad customers at the lowest scores, therefore the cumulative density in defaulted customers shoots up. The red line is the random model (Rand_Model) which indicates that the Bad customers get the similar distribution of scores as the whole sample, in this case the model has no discriminatory power. The blue line is the new developed model (New_Model). The definition of accuracy ratio is given by the following formula:

$$Accuracy\ Ratio\ (AR) = \frac{\alpha}{\alpha + \beta}$$

The larger the AR is, the better the model is. For the random model, AR will be equal to 0. The AR is getting larger when the new model curve shift towards the perfect model's curve.

To facilitate the calculation of AR for VPBank in the future, we have designed an Excel computation tool as follows which has the functions of calculating AR, plotting CAP curve, as well as calculating KS and plotting KS chart for reference purpose. The inputs to the tool are the scores and Good/Bag flag of the customers, please see the "User Manual" for the explanation of how to use the tool.

AR & CAP, KS.xlsx

Worksheet 2: AR & CAP, KS computation tool

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

3. MODEL DEVELOPMENT PROCESS

3.1. DEFINE SAMPLE WINDOW & GOOD/BAD FLAG

Observation window: refers to the time frame within which data are sampled and used in selecting a subset of variables for predicting actual account performance. This is the period of time that applications made are extracted specifically for the scorecard development process.

Generally 12-24 months for account acquisition, 6-12 months for account management

- May be refined by investigating delinquency curves
- The best approach is to design the performance window in equal length for each account, but sample size will be an issue

Performance window: refers to the time frame which data are selected for recording actual account performance. This is the period of time between the observation point and the point where the final performance is assessed.

- Recent enough to be representative of current population
- Dated enough to allow accounts to have established payment behaviors
- Long enough to meet minimum sample size requirements and to accommodate seasonality concerns

3.1.1. OBSERVATION WINDOW, PERFORMANCE WINDOW & GOOD/BAD FLAG

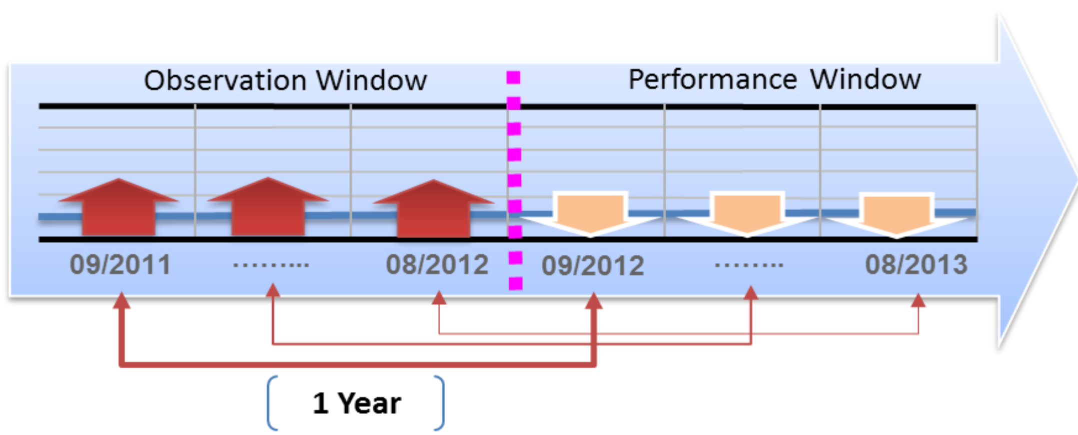


Figure 4: Observation window and performance window

We define observation window from 01/09/2011 to 31/08/2012 and performance window from 01/09/2012 to 31/08/2013. We fix the performance duration for 12 months, which means the Large Corporate scorecard model developed will estimate the probability of default of customers within 12 months.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

3.2. EXTRACT DATA FROM CIC BASED ON OBSERVATION WINDOW

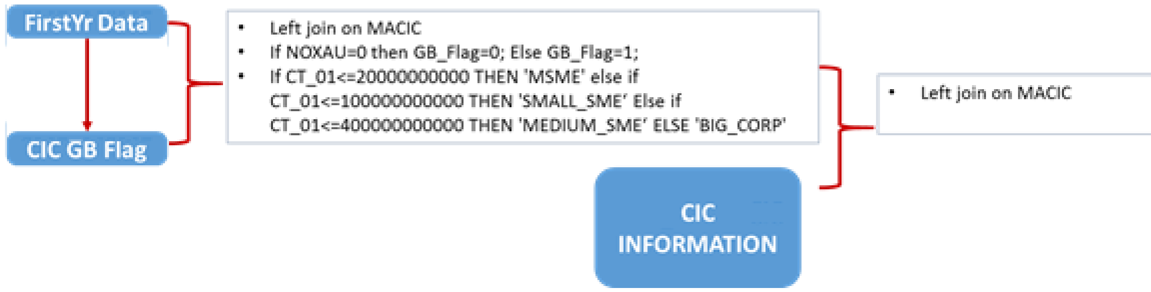


Figure 5: Illustration of extracting CIC data by SQL

The datasets are extracted by SQL code and combined with a few tables in database to generate the final dataset. Please find the below attachments for the datasets:

FirstYr Data.xlsx | CIC GB Flag.xlsx | CIC Info.xlsx

Worksheet 3: First-year data information

Worksheet 4: Good/Bad flag

Worksheet 5: CIC

3.3. DATA CLEANSING

3.3.1. DATA EXCLUSION

After we get the data, it is expected to have some errors such as data entry errors, therefore we need to exclude those incorrect data. We check 30 accounting relationships and exclude those data do not obey the relationships. The structure of our data after exclusion is as follows:

Figure 6: Exclusion of data by accounting relationship checking

Accounting Relationship.txt

Copyright © 2014 EntroFine. All Rights Reserved.

For the details of the accounting relationship rules, please check the text file:

3.3.2. FINANCIAL RATIO CALCULATION

After we extract all the financial/income statement item data from CIC, we calculate all the feasible financial ratios by SQL. We have a list of financial ratio candidates that are used by any other foreign bank as one of the candidates for credit scoring models. For details of financial ratio candidates list, please refer to the following Excel file.

FinancialRatio List v2.0 0114201

Worksheet 6: List of financial ratio candidates

3.3.2.1. RULES DEALING WITH ABNORMAL VALUE WHILE CALCULATING FINANCIAL RATIO

During the financial ratio calculation, there are some cases the denominator is 0, we replace these indeterminate values by either MINIMUM, MAXIMUM or NULL (MISSING) value depending on the nature of the financial ratio and performance of the customers. Please also note that in the final development sample, for those minimum values we assign 8888 to indicate that these are the minimums, for maximum values we assign 9999 and for null (missing) the value to assign is just Null.

For example:

$$DSCR = \frac{Net\ Operating\ Income}{Total\ Debt}$$

Here if we have a customer whose total debt is zero, and if the net operating income is positive (>0), we know this customer has no debts and a positive net operating income will indicate the customer is “Good”. For details of treatment of this abnormal value, please refer to column “Treatment for Abnormal Value” in Worksheet 6.

3.3.3. FINAL DEVELOPMENT DATASET

After completing the above procedure for data extraction and cleansing, we have the following final dataset which contains customer id, financial ratios, CIC information and Good/Bad flag.

FinalDraft_Varselection v13.0.csv

Worksheet 7: Final dataset after data cleansing

3.4. DATA TRANSFORMATION

Before fitting the scorecard model with the data of risk factors (Financial Ratios & CIC information), we need to perform risk factor transformation. Transformation is a mathematical function that transforms the original factor into a new factor which will replace the original factor in the analysis process and will enter into the model if it is selected later on.

The purposes of transformation include:

I. Determine the range of values over which the factor has most impact
II. Smoothing of noise in the data
III. Transform the categorical (qualitative) factors to numeric factors (e.g. WOE transformation)
IV. Replace the missing values by other values (e.g. imputation)
V. Treatment of outliers (e.g. logistic transformation)
VI. Enable comparability of different factors which are measured in different units and scales (e.g. standardisation)

3.4.1. WOE TRANSFORMATION

For those categorical (qualitative) variables (in our dataset, they are CIC9, CIC10, CIC13 (after grouped into 4 categories)), since their values in nature are not measured on the real number line, instead they are measured by categories. We need to transform these factors into numeric variables, to do so, we can calculate the WOEs of these factors and replace the original factors by the WOEs.

The rationale of WOE is some kind of measure reflecting the degree of Good/Bad ratio in each category of a categorical variable. Hence for a categorical variable which has good discriminatory power to Good/Bad, different category will have obviously different Good/Bad ratio and hence significantly different values in WOE. The formula of WOE is

$$WOE = \ln\left(\frac{g_i/G}{b_i/B}\right)$$

, where g_i is the number of "good" obligor of category i, b_i the number of "bad" obligor of category i, G the total number of "good" obligor in sample, B the total number of "bad" obligor in sample

To treat the missing value in calculating WOE, there are two conditions: first, for CIC9, the missing itself is treated as a direct answer to the question: has the borrowing company had overdue debts in any credit institutions in the last 12 months? (DPD more than 90 days), we denote the missing as NA, and it is added on top of Yes and No to represent a single category, in this case the WOE should be calculated for the missing (NA) category; second, for CIC10 and CIC13, we exclude the missing values to calculate the WOE first and will leave the missing values to be addressed by imputation (see Session 3.4.2).

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Table 5, Table 6 and Table 7 illustrate how WOE is calculated for CIC9, CIC10 and CIC13 for Large Corporations.

CIC9 (G= 875, B= 56)	g_i	b_i	g_i/G	b_i/B	Ratio	WOE
YES	0	26	0.0000	0.4643	0.0000	-Inf
NO	346	0	0.3954	0.0000	Inf	Inf
NA	529	30	0.6046	0.5357	1.1285	0.1209

Table 5: WOE calculation for CIC9 of Large Corporations

CIC10 - Geographic factor (G= 875, B= 56)	g_i	b_i	g_i/G	b_i/B	Ratio	WOE
Hanoi	162	19	0.1851	0.3393	0.5457	-0.6057
Duyen hai phia Bac	78	8	0.0891	0.1429	0.6240	-0.4716
Dong Nam Bo	48	1	0.0549	0.0179	3.0720	1.1223
Dong bang song Cuu Long	132	10	0.1509	0.1786	0.8448	-0.1687
Other	36	2	0.0411	0.0357	1.1520	0.1415
Bac Trung Bo	18	0	0.0206	0.0000	Inf	Inf
Mien Trung	45	2	0.0514	0.0357	1.4400	0.3646
Trung du phia Bac	37	5	0.0423	0.0893	0.4736	-0.7474
Dong Ho Chi Minh	319	9	0.3646	0.1607	2.2684	0.8191

Table 6: WOE calculation for CIC10 of Large Corporations

CIC13 (G= 875, B= 56)	g_i	b_i	g_i/G	b_i/B	Ratio	WOE
MISSING	646	15	0.7383	0.2679	2.7563	1.0139
ZERO	79	31	0.0903	0.5536	0.1631	-1.8134
ONE	35	3	0.0400	0.0536	0.7467	-0.2921
MORE THAN ONE	115	7	0.1314	0.1250	1.0514	0.0501

Table 7: WOE calculation for CIC13 of Large Corporations

Note that after the WOE transformation, CIC9 will have values of positive infinity (Inf) and negative infinity (-Inf), these values will cause the use of the WOE of CIC9 to enter the scorecard model become impossible, but in fact CIC9 has strong discriminatory power, instead we will suggest using CIC9 as a knock-out (KO) factor, see Session 3.6.2 for detailed explanation. Similarly, CIC10 has a group (Bac Trung Bo) of no “Bad” customers and so creates a WOE of positive infinity, it cannot enter the model as a risk factor, and we do not suggest it to be used as

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

KO factor, see Session 3.6.2 as well for explanation. For the detailed calculation of the WOE transformation, please see the following Excel file.

WOE.xlsx

Worksheet 8: WOE transformation

3.4.2. IMPUTATION

When a risk factor (including the financial ratios and the CIC risk factors; for CIC9, CIC10 and CIC13, we mean their WOEs) has a missing value, we will do the imputation to replace the missing value by the median of that risk factor. Without the imputation, the whole observation including the Good/Bad flag and the values of other risk factors will not be able to be used in the model fitting process; also the lack of the input value of any risk factor will cause the output score of the scorecard cannot be calculated. Hence imputation is important to preserve the information of the observations in the model development process as well as in the practical implementation.

In theory, there are many methods for imputation. The most common methods are to directly impute the missing values by mean or median of the non-missing values. We prefer using median instead of mean because mean has a problem that it is distorted by the outliers of a highly skewed distribution but median is robust to the outliers, also for a bell-shaped (e.g. normal) or symmetric distribution, the difference between mean and median is little. The idea above is illustrated in the following figure.

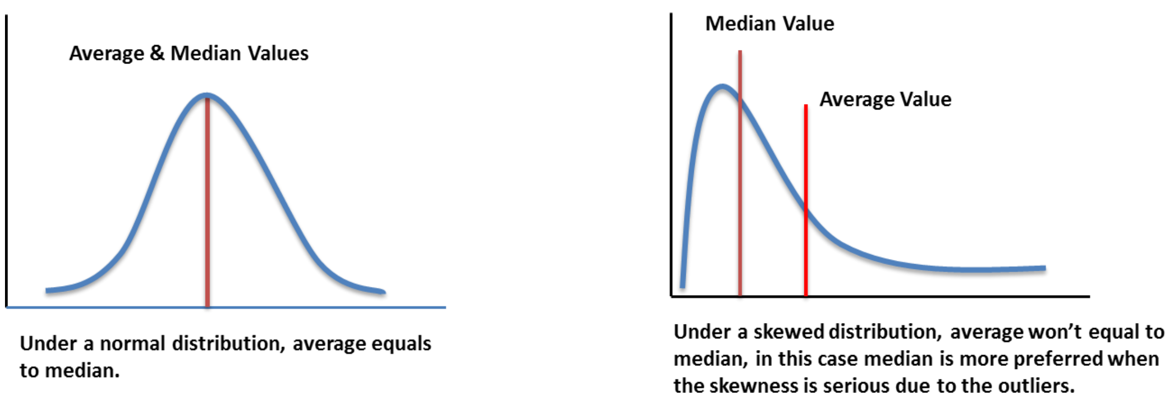


Figure 7: Comparison of mean and median for imputation

Table 8 lists the medians of all risk factors in our sample (WOE for CIC9, CIC10, CIC13), they will be used for imputation when there are missing values of the risk factors.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Category	Variable	Median of Large Corporations	Category	Variable	Median of Large Corporations
Return	Return1	0.0369	DSCR	DSCR1	0.0703
	Return3	0.1508		DSCR2	0.1007
	Return5	0.0451		DSCR3	2.1068
	Return7	0.1859		DSCR4	3.8535
	Return9	0.3679		DSCR5	0.2496
	Return10	0.2871		DSCR6	0.1994
	Return11	0.0993		DSCR7	0.1089
	Return12	0.0801		DSCR8	0.1453
	Return13	0.0403		DSCR9	1.5354
	Return13n	0.0597		DSCR10	0.9772
	Return14	0.1620		DSCR12	0.5124
	Return14n	0.2260		DSCR13	-2.5418
	Return15	0.2817		DSCR14	-90.3456
	Return16	0.2218		DSCR15	-89.8978
	Return17	0.0322		DSCR16	-11.2451
	Return18	0.1208		DSCR17	17.2135
Profitability	Profitability1	0.0910		DSCR18	0.4033
	Profitability2	0.0251		DSCR19	0.0442
	Profitability3	0.0226		DSCR20	0.3807
	Profitability4	0.0200		DSCR21	1.7433
	Profitability5	0.0577	Leverage	Leverage1	0.7124
	Profitability6	0.0439		Leverage2	1.3702
	Profitability13	0.0225		Leverage3	2.4795
	Profitability14	1.0000		Leverage4	2.0348
Efficiency	Efficiency1	1.7104		Leverage5	0.0913
	Efficiency2	59.2647		Leverage6	0.0242
	Efficiency3	46.3662		Leverage7	0.3850
	Efficiency4	7.0235		Leverage8	0.1288
	Efficiency5	10.1965		Leverage9	0.1141
	Efficiency6	9.2828		Leverage10	0.9375
	Efficiency7	3.0823		Leverage18	0.0372
	Efficiency8	113.5242		Leverage19	0.0803
	Efficiency15	29.3786		Leverage20	2.4768
	Efficiency18	0.1424		Leverage21	0.9764

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Efficiency19	6.1984	Size	Size1	27.0914
Liquidity	Liquidity1	1.1380		Size2	27.4652
	Liquidity3	0.1213		Size3	25.1458
	Liquidity4	0.0853		Size4	25.7649
	Liquidity5	1.6801	CIC	CIC1	28.0000
	Liquidity6	0.7232		CIC2	6.0000
	Liquidity11	2.7175		CIC3	40.0000
	Liquidity12	0.0351		CIC4	21.0000
	Liquidity13	0.3218		CIC5	11.0000
	Liquidity16	0.5767		CIC6	0.6685
	Liquidity17	2.3463		CIC7	6.0000
	Liquidity19	1.7756		CIC8	210510.0000
	Liquidity20	1.0761		CIC9 WOE	0.1209
				CIC10 WOE	0.1415
				CIC11	0.0000
				CIC12	0.0000
				CIC13 WOE	1.0139

Table 8: Medians of risk factors for Large Corporations

We also attach the Excel file of the median values of the risk factors for reference.

Median.xlsx

Worksheet 9: Medians of risk factors

3.4.3. LOGISTIC TRANSFORMATION

For the financial ratios (except Profitabililty14 - Profit After Tax, which is a binary variable (equal to 1 or 0), we can use logistic transformation to map the risk factor values to the range of (0,1) but do not distort the ranks of the data. To perform the logistic transformation, we need to first choose the two cutoffs of percentage, e.g. left 4% and right 6%, then the outliers on the “left tail” up to 4-th percentile and the “right tail” from 94-th percentile will be “squeezed” into the left-tailed interval [0,0.04] and right-tailed interval [0.94,1] respectively after the transformation. The logistic transformation will optimize the linearity of the default rate to the risk factor value on the central part of the data (from 4-th percentile to 94-th percentile in our example). The following figure illustrates how the risk factor is transformed and how the outliers are “squeezed” by the logistic transformation.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

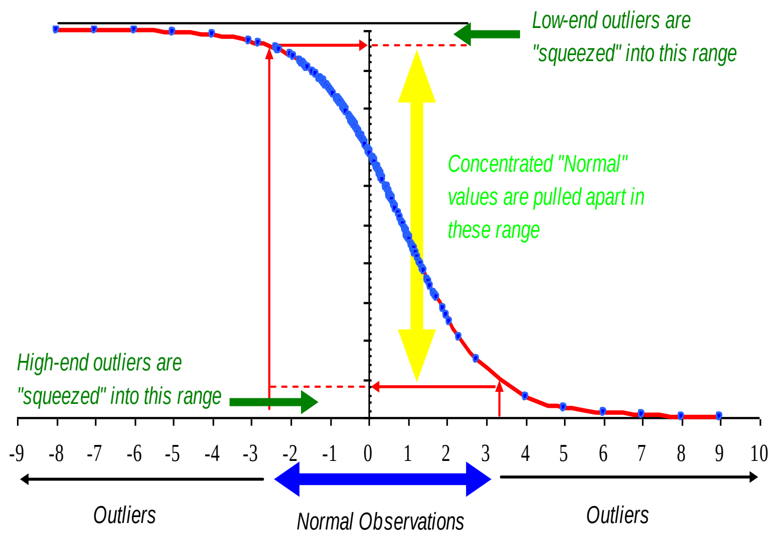


Figure 8 : Illustration of logistic transformation

In mathematical representation, the logistic transformation can be expressed in the following formula with the parameters a and b,

$$X^* = \frac{1}{1+\exp(a+bX)}$$

, where a and b can be solved by the following equations:

$$\alpha_L = \frac{1}{1+\exp\left(a+bF_X^{-1}(\alpha_L)\right)} \qquad 1-\alpha_R = \frac{1}{1+\exp\left(a+bF_X^{-1}(1-\alpha_R)\right)}$$

$$\Downarrow$$

$$a = \frac{F_X^{-1}(1-\alpha_R)\ln\left(\frac{1-\alpha_L}{\alpha_L}\right)-F_X^{-1}(\alpha_L)\ln\left(\frac{\alpha_R}{1-\alpha_R}\right)}{F_X^{-1}(1-\alpha_R)-F_X^{-1}(\alpha_L)} \qquad b = \frac{\ln\left(\frac{\alpha_R}{1-\alpha_R}\right)-\ln\left(\frac{1-\alpha_L}{\alpha_L}\right)}{F_X^{-1}(1-\alpha_R)-F_X^{-1}(\alpha_L)}$$

, where α_L and α_R represent the left and right cutoff percentages respectively, and F_X^{-1} denotes the inverse of empirical distribution function, i.e. $F_X^{-1}(\alpha_L)$ denotes the $100\alpha_L$-th percentile.

To determine α_L and α_R, we can follow the following steps:

1. Choose a pair of values of α_L and α_R (in our analysis, we take the values of 0.02, 0.04, …, 0.2), observe the histograms before and after the transformation, the scatter plot, and PD curve after the logistic transformation, for example, see Figure 9 for Leverage7 of Large Corporations.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

2 . Observe and check whether the scatter plot looks well with no too many serious outliers, the shape of histogram improve from the original with less outliers, and whether the PD curve is smooth, is the monotonicity (i.e. increasing or decreasing linear trend) obvious.

3 . Repeating checking different pairs of values of α_L and α_R, choose the pair which performs as well as possible for the criteria described at step 2 above while keeping the values of α_L and α_R as low as possible.

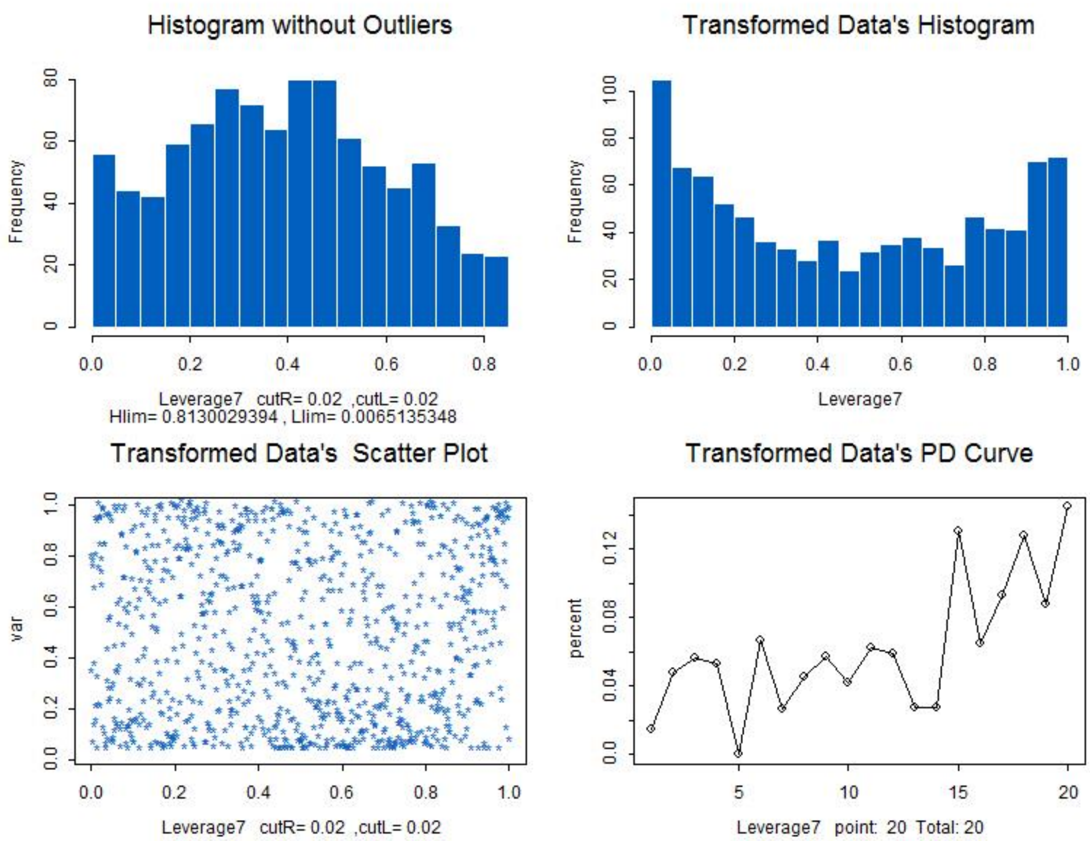


Figure 9: Histograms, scatter plot and PD curve for reference of logistic transformation – Leverage7 of Large Corporations

Table 9 shows the left and right cut off percentages for logistic transformation we chose for all financial ratios for Large Corporations respectively as well as the corresponding parameters a and b.

Category	Variable	CL	CR	a	b	Category	Variable	CL	CR	a	b
Return	Return1	2%	2%	2.5347	-22.4798	**DSCR**	DSCR1	2%	6%	2.4393	-7.8130
	Return3	2%	2%	1.0413	-6.0201		DSCR2	2%	6%	2.8910	-7.2380

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Return5	2%	2%	2.6997	-19.7478
	Return7	2%	2%	1.3593	-5.3485
	Return9	2%	6%	3.4314	-5.8905
	Return10	2%	2%	2.5514	-4.5511
	Return11	2%	2%	3.7362	-20.1327
	Return12	2%	2%	3.0469	-19.6653
	Return13	2%	2%	2.1498	-18.0393
	Return13n	2%	2%	2.4240	-17.2632
	Return14	2%	2%	0.4618	-4.5786
	Return14n	2%	4%	1.7769	-5.1355
	Return15	2%	14%	3.7042	-8.7101
	Return16	2%	6%	3.1119	-7.7626
	Return17	2%	2%	1.7379	-16.3984
	Return18	2%	2%	1.1270	-6.6676
Profitability	Profitability1	2%	2%	3.9438	-15.5745
	Profitability2	2%	4%	2.5152	-23.6197
	Profitability3	2%	6%	2.1511	-27.1599
	Profitability4	2%	4%	2.3489	-27.3512
	Profitability5	2%	2%	3.8051	-16.4967
	Profitability6	2%	12%	2.4983	-31.7813
	Profitability13	2%	6%	2.4257	-22.5208
Efficiency	Efficiency1	2%	12%	4.1381	-1.2745
	Efficiency2	2%	12%	3.9288	-0.0371
	Efficiency3	2%	6%	3.9520	-0.0338
	Efficiency4	2%	18%	4.1787	-0.2795
	Efficiency5	2%	14%	3.8918	-0.0986
	Efficiency6	2%	16%	3.9354	-0.0881
	Efficiency7	2%	12%	4.3048	-0.7049
	Efficiency8	2%	4%	4.1134	-0.0148
	Efficiency15	2%	2%	3.8918	-0.0421
	Efficiency18	2%	10%	3.9316	-15.0374
	Efficiency19	2%	6%	3.9470	-0.1391
Liquidity	Liquidity1	2%	12%	5.8220	-4.0248
	Liquidity3	6%	2%	0.2047	-5.4033
	Liquidity4	2%	2%	1.4595	-9.4867
	DSCR3	2%	8%	3.8110	-0.4483
	DSCR4	2%	2%	3.6554	-0.0002
	DSCR5	2%	16%	3.7527	-6.6016
	DSCR6	2%	16%	2.9439	-6.5714
	DSCR7	2%	16%	2.0326	-6.4692
	DSCR8	2%	14%	2.4662	-5.6391
	DSCR9	2%	12%	4.1523	-1.2031
	DSCR10	2%	2%	4.4671	-3.3112
	DSCR12	2%	18%	2.8058	-2.8182
	DSCR13	14%	2%	-3.9887	-0.8489
	DSCR14	2%	2%	-3.8919	0.0000
	DSCR15	2%	2%	-3.8919	0.0000
	DSCR16	18%	2%	-3.9475	-0.0895
	DSCR17	2%	2%	3.8919	-0.0001
	DSCR18	2%	8%	4.0096	-3.6221
	DSCR19	2%	6%	2.2328	-9.4674
	DSCR20	2%	8%	3.9105	-3.7057
	DSCR21	2%	12%	5.4921	-1.5541
Leverage	Leverage1	2%	2%	5.9767	-10.2047
	Leverage2	2%	6%	3.9064	-0.9322
	Leverage3	2%	4%	4.0046	-0.4392
	Leverage4	2%	8%	4.0076	-0.7086
	Leverage5	2%	14%	3.8918	-6.5352
	Leverage6	2%	4%	3.8918	-15.3994
	Leverage7	2%	2%	3.9547	-9.6513
	Leverage8	2%	8%	3.8918	-3.5073
	Leverage9	2%	10%	3.8918	-10.3286
	Leverage10	2%	6%	3.9133	-1.4558
	Leverage18	2%	2%	3.8918	-13.4595
	Leverage19	2%	12%	3.8918	-6.1334
	Leverage20	2%	4%	3.9398	-0.4697
	Leverage21	2%	2%	17.7905	-21.6824
Size	Size1	2%	2%	38.2681	-1.3968
	Size2	2%	10%	82.4992	-2.9398
	Size3	2%	4%	30.3669	-1.2208

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Liquidity5	2%	6%	5.6515	-1.6677		Size4	2%	2%	29.4587	-1.1438
	Liquidity6	2%	2%	5.1618	-9.1107						
	Liquidity11	2%	16%	4.3635	-1.0902						
	Liquidity12	2%	6%	0.4427	-8.0207						
	Liquidity13	4%	2%	-0.3333	-2.8605						
	Liquidity16	2%	10%	4.4714	-6.1161						
	Liquidity17	2%	18%	4.6063	-0.8575						
	Liquidity19	2%	18%	4.4955	-1.1478						
	Liquidity20	2%	10%	5.7303	-4.2563						

Table 9: Logistic Transformation left and right cutoffs for Large Corporations

The following Excel file shows the detailed calculation for the parameters a and b in the logistic transformation.

Logistic Transform.xlsx

Worksheet 10: Logistic transformation

3.4.4. STANDARDISATION

After the risk factors finish the above transformations, they will enter the last step of transformation – standardisation. Traditionally in statistics, standardisation means to minus a variable by its sample mean and then divide it by its sample standard deviation. In our analysis we will further multiply the resulting factor by 50, i.e. $X^* = 50 \times (X - \bar{X})/S$. There are mainly two purposes of our standardisation, first is to ensure the comparability of different risk factors in the same scale which are originally measured in different units and scales, so that the regression coefficients in the logistic regression model will have comparability and hence making the weights of the coefficients to represent the relative dominance of the risk factors in the model become possible; second is to preserve the same comparability in scale throughout the whole scoring process, so that the final score is also measured in the same standardised scale, which improves the intuitiveness and interpretation of the final score, and also preserve the comparability of the final score to the sub-category score.

The rationale of the multiplier 50 comes from the fact that for a standard normally distributed random variable, approximately 95% of chance it will lie between -2 and +2. When it is multiplied by 50, 95% of chance we will get the values from -100 to 100. Since the values of -2 and +2 are less intuitive than -100 and +100, also the values falling in the range (-2,2) are too dense, not easy for distinguishing the difference (we need to read the values at 2 to 3 d.p.), the multiplier 50 can enlarge the scale to -100 and 100 so improve both the intuitiveness and the comparability of

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

the values. Hence taking normal distribution as a reference, we can treat those values falling out -100 and 100 as the outliers. A direct benefit to the interpretation of the final score is that we can treat 0 as the average level of the score, so positive score means better than average, a score higher than 100 will be an exceptionally good customer, for negative score, vice versa.

Table 10 presents the sample mean and sample SD of the risk factors (after all the transformations described in previous sub-sessions) in our sample for the use in standardisation.

Category	Variable	Mean	SD	Category	Variable	Mean	SD
Return	Return1	0.2615	0.2482	DSCR	DSCR1	0.2551	0.2704
	Return3	0.4945	0.2108		DSCR2	0.2304	0.2713
	Return5	0.2528	0.2523		DSCR3	0.1778	0.2816
	Return7	0.4465	0.2201		DSCR4	0.2774	0.4212
	Return9	0.3335	0.2901		DSCR5	0.2945	0.3410
	Return10	0.3074	0.2395		DSCR6	0.3320	0.3319
	Return11	0.2550	0.2582		DSCR7	0.3626	0.3161
	Return12	0.2792	0.2552		DSCR8	0.3202	0.3166
	Return13	0.2943	0.2453		DSCR9	0.2474	0.3169
	Return13n	0.3011	0.2554		DSCR10	0.2811	0.2299
	Return14	0.5749	0.1997		DSCR12	0.3621	0.3321
	Return14n	0.4130	0.2477		DSCR13	0.6905	0.3353
	Return15	0.3554	0.3247		DSCR14	0.7056	0.4339
	Return16	0.3135	0.2867		DSCR15	0.7056	0.4339
	Return17	0.3230	0.2414		DSCR16	0.7378	0.3677
	Return18	0.4675	0.2299		DSCR17	0.2969	0.4346
Profitability	Profitability1	0.1761	0.2346		DSCR18	0.2172	0.2958
	Profitability2	0.2350	0.2469		DSCR19	0.2561	0.2619
	Profitability3	0.2919	0.2638		DSCR20	0.2223	0.2971
	Profitability4	0.2500	0.2448		DSCR21	0.2276	0.3291
	Profitability5	0.1326	0.2102	Leverage	Leverage1	0.6592	0.3074
	Profitability6	0.3724	0.3034		Leverage2	0.1916	0.2720
	Profitability13	0.2441	0.2588		Leverage3	0.1518	0.2454
	Profitability14	0.9441	0.2298		Leverage4	0.2124	0.2928
Efficiency	Efficiency1	0.2926	0.3286		Leverage5	0.2273	0.3458
	Efficiency2	0.3107	0.3287		Leverage6	0.1550	0.2672
	Efficiency3	0.2256	0.2908		Leverage7	0.4693	0.3317
	Efficiency4	0.2956	0.3557		Leverage8	0.1689	0.2943

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Efficiency5	0.2468	0.3463		Leverage9	0.2518	0.3320
	Efficiency6	0.2413	0.3579		Leverage10	0.2031	0.2770
	Efficiency7	0.2754	0.3227		Leverage18	0.1571	0.2641
	Efficiency8	0.1995	0.2572		Leverage19	0.2090	0.3311
	Efficiency15	0.1798	0.2504		Leverage20	0.1661	0.2494
	Efficiency18	0.2949	0.3153		Leverage21	0.8226	0.2786
	Efficiency19	0.1633	0.2711	**Size**	Size1	0.4303	0.2831
Liquidity	Liquidity1	0.3425	0.2958		Size2	0.3045	0.3205
	Liquidity3	0.5911	0.2606		Size3	0.5641	0.2662
	Liquidity4	0.4088	0.2763		Size4	0.5013	0.2825
	Liquidity5	0.1938	0.2879	**CIC**	CIC1	28.5951	13.2252
	Liquidity6	0.6699	0.3137		CIC2	9.2095	10.0603
	Liquidity11	0.3525	0.3374		CIC3	48.5317	36.6961
	Liquidity12	0.5003	0.2215		CIC4	25.8110	20.1299
	Liquidity13	0.6963	0.2576		CIC5	13.2997	10.5685
	Liquidity16	0.3787	0.3169		CIC6	1.5354	3.0487
	Liquidity17	0.2766	0.3638		CIC7	7.0827	5.0982
	Liquidity19	0.2800	0.3576		CIC8	449865.0849	969413.0616
	Liquidity20	0.3434	0.2876		CIC9 WOE	NA	NA
					CIC10 WOE	Inf	NA
					CIC11	0.0838	0.2772
					CIC12	0.5682	1.3867
					CIC13 WOE	0.3979	0.8973

Table 10: Means and SDs of transformed risk factors for Large Corporations

The sample means and sample SDs of the above tables are summarised in the following Excel file for reference.

Mean & SD.xlsx

Worksheet 11: Means and SDs of transformed risk factors

3.5. SINGLE FACTOR ANALYSIS

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

After the data transformation, the risk factors now are ready for use to build the scorecard. We will first do the single factor analysis to the financial ratios (CIC factors will be analysed in Session 3.6). The aim is to examine each financial ratio one-by-one, to preliminarily short-list a list of candidates for further work of selection of risk factors. If a financial ratio passes all the following criteria, then it will be short-listed:

1. Absolute value of AR (Accuracy Ratio) > 11%
2. Sign of observed AR is consistent to the theoretical sign (theoretical trend of the default rate to the risk factor)
3. P-value of beta (regression coefficient of the single-factor logistic regression) < 0.1
4. Percentage of abnormal value (before abnormal value treatment) < 15%
5. Percentage of missing of the original risk factor (before imputation) < 20%

Next we will explain the above criteria one by one in details:

1. Absolute value of AR (Accuracy Ratio) > 11%

   AR, as we defined in Session 2.4, is a statistical measure of the discriminatory power of the score of a scorecard to Good/Bad flag. In general, AR can be applied to any single risk factor to measure the risk factor's discriminatory power. The value of AR ranges from -1 to 1. When AR=0, it means the risk factor has no discriminatory power to distinguish Good/Bad, there is no difference to randomly assign Good or Bad to any given value of the risk factor. When AR>0, it means there is a tendency in higher chance for a higher value of the risk factor to get a "Good" than a "Bad" customer. When AR<0, the higher the value of the risk factor, the more likely a customer is a "Bad" customer than a "Good" one.

   To ensure a risk factor has high enough discriminatory power to Good/Bad, AR must be high enough in magnitude either on the positive side or on the negative side. As a rule of thumb, we set the rule the absolute value of AR must be greater than 11%.

2. Sign of observed AR is consistent to the theoretical sign (theoretical trend of the default rate to the risk factor)

   In addition to checking the magnitude of AR to ensure large enough discriminatory power, we also need to check whether the sign (positive or negative) of AR is consistent to the theoretical direction. For example, based on the common business sense and theory, and also our experience, when the financial ratios of the category "Leverage" increase, it should be more likely for a company to default, hence reducing the chance of being "Good", and the theoretical sign of AR should be negative; when the financial ratios of the category "Liquidity" increase, it should be less likely for a company to default, and so a higher chance to be "Good", hence the theoretical sign of AR should be positive.

   When the observed sign of AR is inconsistent to the theoretical sign, we will forgo that financial ratio, otherwise it is possible to produce misleading score in the final scorecard, and harmful to the business implementation.

3. P-value of beta (regression coefficient of the single-factor logistic regression) < 0.1

Apart from AR, we will also test the linear dependence of the risk factor to the default rate in a single-factor logistic regression model, i.e. for a given risk factor x, and the probability of default p, the model is given by

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

Then we will test the following hypothesis,

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

To access the significance of the test, we will consider the p-value. Instead of using the traditional rule of thumb of p-value 0.05, we use a looser p-value cutoff 0.1, the reason is that we are at the preliminary stage of selecting discriminating risk factor, the rule should not be too strict to exclude too many potential risk factors, 0.1 is a feasible level which allows more potential risk factors to be examined while will not be too loose to unnecessarily shortlist too many candidates.

4. Percentage of abnormal value (before abnormal value treatment) < 15%

   Although we have special treatment to abnormal value in the model, the score obtained from a risk factor with abnormal value will be less accurate than the normal value. Hence, we should avoid using those risk factors which have too high chance to get an abnormal value, it will affect the quality of the scorecard in application. So we set the rule to exclude those risk factors which have at least 15% of abnormal value.

5. Percentage of missing of the original risk factor (before imputation) < 20%

   Apart from abnormal value, missing value is also a problem which will affect the scorecard quality. Since the missing value of a risk factor will be imputed by the median (the exception is CIC9, the missing is treated as a single category NA, and we use WOE to replace the categories), it will on one hand affect the fit and accuracy of the model in the development stage, on the other hand a too high chance of missing will make the data availability problematic, and the imputed value will affect the accuracy of the score in application. Those risk factors with at least 20% of missing value will so be excluded from the shortlist.

Table 11 presents the results of the single-factor analysis for the assessment of the rules listed above. Note that the highlighted financial ratios are the shortlisted candidates which pass all the criteria and hence they will form the list of candidates for the multi-factor analysis in Session 3.7.

Category	Variable	Beta	P-value	AR	Theoretical Sign	Consistency	% Abnormal	% Missing
Return	Return1	-0.0178	0.00	34.92%	+	Y	0.0%	0.0%
	Return3	-0.0137	0.00	29.62%	+	Y	0.9%	0.0%
	Return5	-0.0183	0.00	35.48%	+	Y	0.0%	0.0%
	Return7	-0.0144	0.00	30.75%	+	Y	0.9%	0.0%
	Return9	-0.0037	0.22	8.34%	+	Y	0.9%	0.0%
	Return10	-0.0050	0.12	11.57%	+	Y	0.9%	0.0%

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Return11	-0.0135	0.00	28.81%	+	Y	0.0%	0.0%
	Return12	-0.0122	0.00	26.97%	+	Y	0.0%	0.0%
	Return13	-0.0153	0.00	30.60%	+	Y	0.0%	0.0%
	Return13n	-0.0162	0.00	33.94%	+	Y	0.0%	0.0%
	Return14	-0.0104	0.00	25.51%	+	Y	0.9%	0.0%
	Return14n	-0.0086	0.01	22.96%	+	Y	0.1%	0.0%
	Return15	-0.0057	0.07	18.65%	+	Y	0.4%	0.0%
	Return16	-0.0057	0.08	17.17%	+	Y	0.4%	0.0%
	Return17	-0.0115	0.01	21.63%	+	Y	0.0%	0.0%
	Return18	-0.0118	0.00	28.40%	+	Y	0.4%	0.0%
Profitability	Profitability1	-0.0050	0.16	6.59%	+	Y	0.0%	0.0%
	Profitability2	-0.0112	0.02	20.80%	+	Y	0.0%	0.0%
	Profitability3	-0.0100	0.01	19.18%	+	Y	0.0%	0.0%
	Profitability4	-0.0109	0.02	20.13%	+	Y	0.0%	0.0%
	Profitability5	-0.0050	0.20	2.21%	+	Y	0.0%	0.0%
	Profitability6	-0.0034	0.26	9.65%	+	Y	0.0%	0.0%
	Profitability13	-0.0117	0.01	19.98%	+	Y	0.2%	0.0%
	Profitability14	-0.0067	0.00	13.06%	+	Y	0.0%	0.0%
Efficiency	Efficiency1	-0.0119	0.00	29.31%	+	Y	0.0%	0.0%
	Efficiency2	0.0093	0.00	-29.06%	-	Y	0.2%	0.0%
	Efficiency3	0.0067	0.00	-26.00%	-	Y	0.0%	0.0%
	Efficiency4	-0.0092	0.01	31.04%	+	Y	0.5%	0.0%
	Efficiency5	-0.0008	0.79	4.37%	+	Y	23.3%	0.0%
	Efficiency6	-0.0058	0.09	21.69%	+	Y	0.0%	0.6%
	Efficiency7	-0.0086	0.02	30.99%	+	Y	0.0%	0.0%
	Efficiency8	0.0070	0.00	-32.24%	-	Y	0.0%	0.2%
	Efficiency15	0.0061	0.01	-11.84%	+	N	0.0%	0.2%
	Efficiency18	0.0100	0.00	-31.04%	+	N	0.0%	0.0%
	Efficiency19	-0.0049	0.17	11.67%	+	Y	0.9%	0.0%
Liquidity	Liquidity1	-0.0078	0.03	20.32%	+	Y	0.0%	0.0%
	Liquidity3	-0.0061	0.02	20.32%	+	Y	0.0%	0.0%
	Liquidity4	-0.0097	0.00	23.14%	+	Y	0.0%	0.0%
	Liquidity5	0.0009	0.73	9.60%	+	Y	0.0%	0.0%
	Liquidity6	-0.0028	0.28	9.96%	+	Y	0.0%	0.0%
	Liquidity11	-0.0098	0.01	27.31%	+	Y	0.0%	0.0%
	Liquidity12	-0.0056	0.05	12.42%	+	Y	0.0%	0.0%

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Liquidity13	-0.0058	0.02	19.58%	+	Y	0.9%	0.0%
	Liquidity16	-0.0055	0.07	17.87%	+	Y	0.0%	0.0%
	Liquidity17	-0.0065	0.06	26.71%	+	Y	4.5%	0.0%
	Liquidity19	-0.0054	0.10	18.87%	+	Y	0.6%	0.0%
	Liquidity20	-0.0068	0.04	17.51%	+	Y	0.0%	0.0%
DSCR	DSCR1	-0.0092	0.03	28.01%	+	Y	0.0%	0.0%
	DSCR2	-0.0098	0.02	31.53%	+	Y	0.0%	0.0%
	DSCR3	-0.0081	0.06	33.10%	+	Y	0.5%	0.0%
	DSCR4	-0.0059	0.07	30.06%	+	Y	29.1%	0.0%
	DSCR5	-0.0177	0.00	39.11%	+	Y	1.4%	0.0%
	DSCR6	-0.0164	0.00	34.71%	+	Y	1.4%	0.0%
	DSCR7	-0.0164	0.00	33.03%	+	Y	1.4%	0.0%
	DSCR8	-0.0193	0.00	37.98%	+	Y	1.2%	0.0%
	DSCR9	-0.0131	0.01	30.62%	+	Y	1.2%	0.0%
	DSCR10	-0.0118	0.00	24.61%	+	Y	0.0%	0.0%
	DSCR12	-0.0162	0.00	33.74%	+	Y	1.4%	0.0%
	DSCR13	0.0090	0.01	-27.50%	+	N	0.0%	0.0%
	DSCR14	0.0037	0.22	-25.79%	+	N	29.1%	0.0%
	DSCR15	0.0037	0.22	-25.80%	+	N	29.1%	0.0%
	DSCR16	0.0030	0.32	-10.10%	+	N	3.5%	0.0%
	DSCR17	-0.0049	0.12	28.10%	+	Y	29.1%	0.0%
	DSCR18	-0.0099	0.02	27.89%	+	Y	0.0%	0.0%
	DSCR19	-0.0109	0.02	22.13%	+	Y	0.0%	0.0%
	DSCR20	-0.0100	0.02	26.86%	+	Y	0.0%	0.0%
	DSCR21	-0.0129	0.01	33.56%	+	Y	1.4%	0.0%
Leverage	Leverage1	0.0113	0.00	-27.68%	-	Y	0.0%	0.0%
	Leverage2	0.0064	0.00	-34.53%	-	Y	0.9%	0.0%
	Leverage3	0.0053	0.01	-27.86%	-	Y	0.9%	0.0%
	Leverage4	0.0053	0.02	-21.13%	-	Y	0.9%	0.0%
	Leverage5	0.0068	0.00	-26.43%	-	Y	0.9%	0.0%
	Leverage6	0.0062	0.00	-24.18%	-	Y	0.0%	0.0%
	Leverage7	0.0115	0.00	-32.42%	-	Y	0.0%	0.0%
	Leverage8	0.0062	0.00	-23.51%	-	Y	0.9%	0.0%
	Leverage9	0.0075	0.00	-23.54%	-	Y	0.4%	0.0%
	Leverage10	0.0054	0.02	-25.54%	-	Y	0.4%	0.0%

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

	Leverage18	0.0059	0.01	-20.74%	-	Y	0.0%	0.0%
	Leverage19	0.0061	0.01	-22.41%	-	Y	0.1%	0.0%
	Leverage20	0.0053	0.01	-23.38%	-	Y	0.1%	0.0%
	Leverage21	-0.0062	0.01	24.18%	+	Y	0.0%	0.0%
Size	Size1	0.0069	0.01	-22.36%	+	N	0.0%	0.0%
	Size2	-0.0013	0.65	-4.77%	+	N	0.0%	0.0%
	Size3	-0.0032	0.24	9.43%	+	Y	0.0%	0.0%
	Size4	0.0009	0.75	-2.47%	+	N	0.0%	0.0%

Table 11: Results of single-factor analysis for SME

The following spreadsheet is the working file we used for the single factor analysis, it as well contains other useful summarising information and statistics for the financial ratios.

SFA – Large Corporate.xlsx

Worksheet 12 : Single-factor analysis for Large Corporations

3.6. CIC FACTORS ANALYSIS

We base on the single factor analysis described in the previous session to shortlist the candidates of financial ratios, these financial ratios will be further selected by the multi-factor analysis (see Session 3.7). However, for the CIC factors, we will directly select and determine the CIC factors which will directly enter the scorecard model during the multi-factor analysis process.

3.6.1. ACCURACY RATIO & PERCENTAGE OF MISSING

Table 12 below summarises the analysis result for CIC information. The preliminary rule to shortlist the CIC risk factors are to keep only those have Accuracy Ratio (AR) > 11% and percentage of missing < 20%. Among the shortlisted CIC risk factors (i.e. CIC2, CIC7 – CIC13), we will further do some correlation analysis to remove those redundant risk factors which have high correlation with the other financial ratios or CIC risk factors (high correlation between two risk factors mean they have overlapping of similar information).

After the analysis, CIC8 and CIC13 are removed due to the high correlations, CIC9, CIC11 and CIC12 will be proposed to be the knock-out (KO) factors (i.e. the factors which provide the decision rules without scoring to directly reject the customers fail to meet the rules). CIC10 is removed because it has an infinite value of WOE on one group (i.e. Bac Trung Bo), which makes CIC10 being a risk factor in the logistic regression model impossible, also the decision rule to be

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

set to make CIC10 a KO factor is unclear and inappropriate with business consideration. Although CIC2 performs not badly in the Large Corporate sample, we recorded over 20% missing rate in the SME/MSMEs sample, and taking into account the small sample size of Large Corporations, for more conservativeness in the statistical sense and security of data availability, we also do not recommend CIC2. Hence finally only CIC7 is selected for the risk factor to enter the scorecard model (in the multi-factor analysis in Session 3.7, CIC7 will be the compulsory risk factor).

Variable	Definition	AR	% Missing	KO Factor?	Selected?
CIC1	Months of oldest bank contract	-7.49%	12.46%	N	N
CIC2	Month of youngest bank contract	31.29%	12.46%	N	N
CIC3	Inquiry times in last 12 months	4.48%	0.00%	N	N
CIC4	Inquiry times in last 6months	10.00%	0.43%	N	N
CIC5	Inquiry times in last 3months	8.88%	2.58%	N	N
CIC6	Q1: When was the borrowing company founded? (day/month/year)	-3.01%	2.58%	N	N
CIC7	Q16: What is the number of lenders outstanding from the CIC form?	-35.32%	0.00%	N	Y
CIC8	Q17: How much is the company's total outstanding lending balance from CIC form?	-35.27%	0.00%	N	N
CIC9	Q18: Has the borrowing company had overdue debts in any credit institutions in the last 12 months? (DPD more than 90 days) YES/NO/NA	67.61%	0.00%	Y	N
CIC10	Geography Information	34.14%	0.00%	N	N
CIC11	Has the customer at observation point has debts in group 2 or not?	-44.29%	0.00%	Y	N
CIC12	How many times a customer have debts in group 2 in the past 12 months from observation point?	-55.82%	0.00%	Y	N
CIC13	How many months since the latest debts in group 2 the customer have?	55.12%	0.00%	N	N

The highlighted CIC factors has AR>11% and percentage of missing <20%

Table 12: List of CIC risk factors

3.6.2. CORRELATION ANALYSIS & KNOCK-OUT FACTORS

We have examined the correlations between CIC7 – CIC13 and the correlations to the other financial ratios. After the analysis, CIC8 is removed due to its highly positive Pearson correlation with CIC7, and CIC13 is removed due to its highly negative Pearson correlations to CIC11 and CIC12, see the following table.

	CIC7

	CIC11	CIC12

CIC8	0.54

CIC13	-0.75	-0.70

Table 13: Pearson correlations of CIC8 to CIC7 and that of CIC13 to CIC11 and CIC12

In addition to the consideration of correlation, CIC 11 and CIC 12 have high absolute values of AR, hence we propose them to be the KO factors. Moreover, CIC9 has very high value of AR (>65%), so it will directly be used as the KO factor as well. Table 14 summarizes the CIC KO factors with their decision rules to "knock-out" the customers who fail to meet the rules.

Variable	Definition	Rule of Knock-out
CIC9	Q18: Has the borrowing company had overdue debts in any credit institutions in the last 12 months? (DPD more than 90 days) YES/NO/NA	If YES, Reject.
CIC11	Has the customer at observation point has debts in group 2 or not?	If YES, Reject.
CIC12	How many times a customer have debts in group 2 in the past 12 months from observation point?	If number of times >2, Reject.

Table 14: Summary of CIC knock-out (KO) factors

3.6.3. SELECTED CIC RISK FACTOR – CIC7 (NUMBER OF LENDERS)

The remaining CIC factors which are not excluded and not chosen for KO factors are CIC2, CIC7 and CIC 10. As we explained in Session 3.6.1, CIC2 is not suggested for security of data availability as it was recorded over 20% of missing values in SME/MSMEs. We will visualize the discriminatory power of CIC7 and CIC10 by plotting the default rate curves to see the trend. We start from CIC10 first and describe the problem of using it in the scorecard, then we will move to CIC7.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

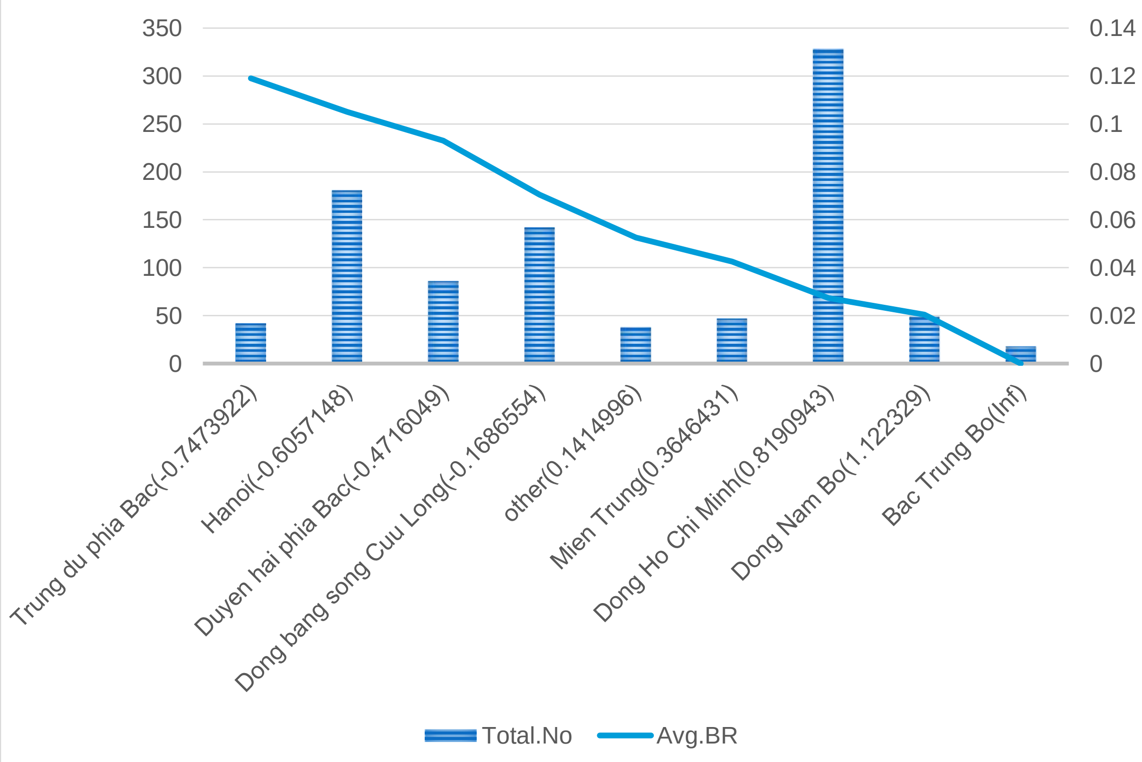


Figure 10: Plot of frequency, default rate against CIC10 (Geography Information) with WOE for Large Corporations

Figure 10 plots the default rate against CIC10 (Geography Information) from the region with highest default rate (i.e. Trung du phia Bac) to the region with lowest default rate (i.e. Bac Trung Bo) for Large Corporations. Note that we also include the values of WOE in the parentheses next to the region names for reference, it is clear from left to right the WOE is in ascending order. The blue bars denote the frequency of observations in the corresponding regions.

Since Bac Trung Bo gets a WOE of positive infinity (Inf), the use of CIC10 as a risk factor of the scorecard model is impossible, because in this case the regression coefficient cannot be estimated. Moreover, the use of CIC10 as a KO factor is also infeasible. Firstly, the AR of CIC10 is 34.14%, it is just merely strong enough to be used as a KO factor. Secondly, and this is the key reason, the region “Hanoi”, has the second highest default rate and so the second lowest WOE, if we define the knock-out rule by setting cutoff of WOE, then any cutoff greater than the WOE of Hanoi would reject all the customers from Hanoi, this will be a conflict of interest to VPBank’s business goal. If other knock-out rules are considered, then some regions have to be decided to be knocked-out, the reasons of choosing these regions will be difficult to be justified, not solely on the statistical consideration of default rates, but also from the business views. Rejecting any regions would restrict the volume of business and it may unnecessarily preventing too much risk, finally leading to lower profitability at a comparable risk level. Based on these considerations, CIC10 is not recommended to be used as a KO factor as well.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

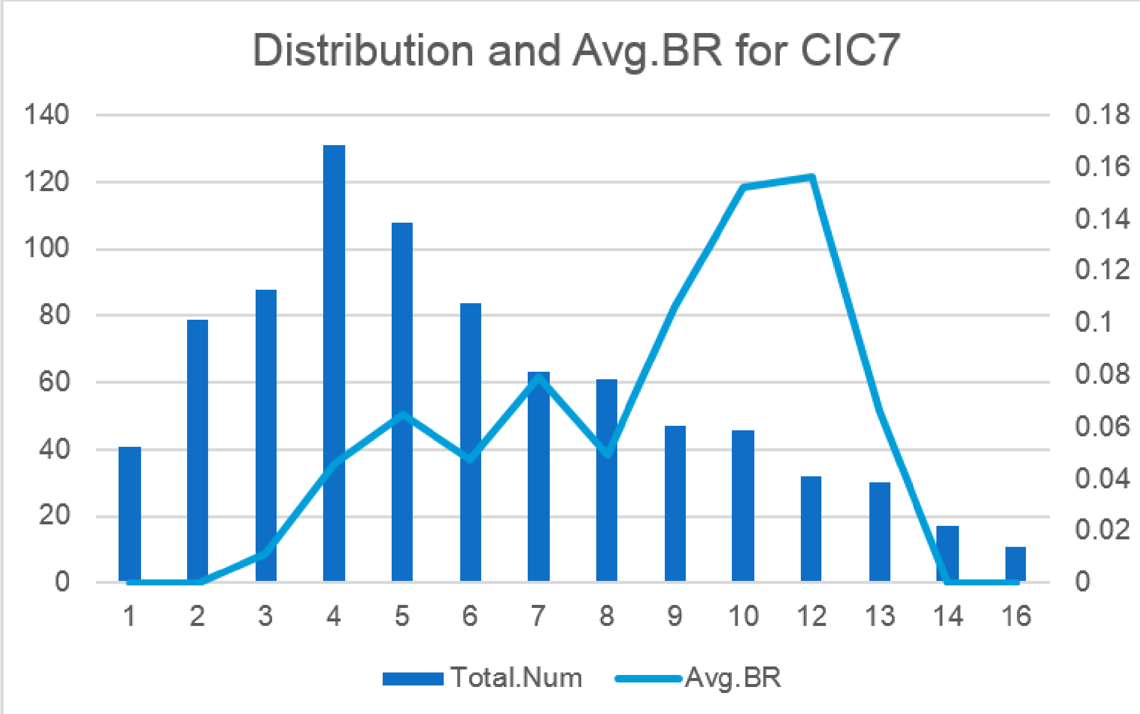


Figure 11: Plot of Frequency, Default Rate against CIC7 for Large Corporations

Figure 11 is the plot of default rate and frequency against CIC7 (Q16: What is the number of lenders outstanding from the CIC form?) for Large Corporations. We can observe a quite clear increasing tread in default rate for the number of lenders ≤ 12. For number of lenders ≥ 12, the trend is decreasing but since the frequencies are relatively low, the default rate estimates are relatively unstable and also they would not affect the overall increasing trend too much.

Since CIC7 has quite good AR of -35.32%, and the linear trend from the graph is quite clear, we will include it into our scorecard model. In the next session of multi-factor analysis, CIC7 will appear as a compulsory risk factor in the model, and we will select one financial ratio from each category of Return, Profitability, Efficiency, Liquidity, DSCR and Leverage (Size is not considered as no one of the risk factors of Size pass the shortlisting by the single factor analysis in Session 3.5) to form the final scorecard model.

3.7. MULTI-FACTOR ANALYSIS

After we gain the shortlisted candidates from the single factor analysis, then we will do the multi-factor analysis to get our final selected risk factors for our scorecard model.

The method we adopt to build the scorecard is the multi-factor logistic regression which is the logistic regression of fitting the Good/Bad flag against the risk factors selected. We will define the score and the weight of each risk factor from the logistic regression model, so that we can

Copyright © 2014 EntroFine. All Rights Reserved.

calculate the Accuracy Ratio (AR) of the score obtained from the model and it will represent the performance (discriminatory power) of the scorecard.

Here comes a question. How do we select a good enough combination of risk factors from our list of candidates to form our final scorecard model? Our answer is to use the procedure of best-subset selection, which is a commonly used method in the model selection problem in Statistics, it is powerful to help achieving an optimal model based on the pre-specified criteria.

3.7.1. WEIGHTS OF RISK FACTORS & SCORE OF THE SCORECARD

Suppose we have a combination of a given set of risk factors, we can build the multi-factor logistic regression model as follows:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

, where k is the number of risk factors and each x_i is a risk factor selected in the model

From the model we can define the weight of each risk factor (the risk factors must be standardised before fitting the model so that each weight is valid to represent the relatively dominance of each risk factor):

$$w_i = \frac{-\beta_i}{\sum_{i=1}^{k} |\beta_i|}$$

Hence from the weights defined above, the score can be expressed as a weighted sum of the risk factors:

$$score = \sum_{i=1}^{k} w_i x_i$$

Then this score will be used for the prediction of Good/Bad for a customer, and we can treat this scoring formula as the main component of the scorecard model. The higher is the score, the more likely is "Good" (less chance to default) a customer the model predicts. So we can calculate the Accuracy Ratio (AR) of the score vector we obtained for all customers in the sample, and the AR will measure the discriminatory power of the score and so the prediction performance of the scorecard model.

Note that the intercept disappears from the scoring formula because removing the intercept will not affect the ranking of the scores of the observations, by choosing an appropriate cutoff of the score will help making the good prediction of Good/Bad without the need of intercept in the model. Moreover, since the risk factors are all standardised with the multiplier 50, and the score is a weighted combination of the risk factors, we can preserve the interpretation what the score at 0 represent the mean level of score in the sample, and taking the normal distribution as a reference, the scores lie beneath -100 and 100 can be treated as outliers. One benefit of removing the intercept in the scoring function is to preserve this interpretation of score without distorting the mean level at 0 and the referencing central interval of score from -100 to 100.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

3.7.2. BEST-SUBSET SELECTION

Then we will explain how we use the best-subset selection to choose our combination of risk factors for the scorecard.

In Statistics, best-subset selection is a model selection method which tries to examine all possible combinations of risk factors - to each combination of risk factors, the model will be fit to that combination, and a model performance measure will be evaluated to the model (AR is used in our case), then we can sort all models according to that model performance measure, so that we can get the best combinations of risk factors in terms of that performance measure.

Our procedure of best-subset selection is a modified procedure to which traditionally means, where the usual procedure in Statistics is to test all possible combinations of risk factors, but in our procedure we restrict the scope so that in each iteration, we choose only one risk factor from each category of Return, Profitability, Efficiency, Liquidity, DSCR, Leverage (Size is not considered as no one of the risk factors of Size pass the shortlisting by the single factor analysis in Session 3.5), then combine with the CIC factor we selected, i.e. CIC7, to run the multi-factor logistic regression, then we will calculate the AR of the score obtained from that multi-factor logistic regression model. In addition, we add two exclusion conditions so that when one of the following condition is met, the combination of that iteration will be forgone:

1. any one Pearson’s correlation between two risk factors is greater than 0.5 or smaller than -0.5
2. the sign of regression coefficient of any risk factor in the multi-factor logistic regression model is different to the sign of regression coefficient in the single factor logistic regression of that risk factor

In each iteration, if it does not meet the above exclusion conditions, we will save the combination of risk factors chosen of that iteration, the weight of each risk factor, and the AR of the scores. The above process is run until all possible combinations of risk factors based on our method are examined, then the resulting combinations are sorted from the highest AR to the lowest AR, the models with the highest AR will be considered for the candidate of the final scorecard model.

The following table lists the top 20 scorecard models in terms of AR of the score for Large Corporations. Note that we restrict to only those models have at least 5% in the absolute value of each weight to avoid too low weight in any category.

Rank	AR	Return	Profitability	Efficiency	Liquidity	DSCR	Leverage	CIC
1	51.13%	Return14n 14.05%	Profitability13 18.31%	Efficiency3 -19.33%	Liquidity4 7.06%	DSCR9 15.09%	Leverage10 -11.57%	CIC7 -14.59%
2	51.07%	Return14n 11.57%	Profitability13 20.22%	Efficiency3 -18.4%	Liquidity4 5.35%	DSCR9 19.46%	Leverage20 -10.51%	CIC7 -14.49%
3	51.05%	Return14n 13.12%	Profitability13 15.99%	Efficiency3 -18.52%	Liquidity16 8%	DSCR9 19.46%	Leverage4 -11.65%	CIC7 -13.26%
4	50.96%	Return14n 9.62%	Profitability13 21.49%	Efficiency8 -19.66%	Liquidity4 6.18%	DSCR9 19.76%	Leverage20 -9.53%	CIC7 -13.76%
5	50.75%	Return14n 13.12%	Profitability13 17.33%	Efficiency3 -21.09%	Liquidity16 9.85%	DSCR9 14.15%	Leverage10 -10.93%	CIC7 -13.54%
6	50.74%	Return14n	Profitability2	Efficiency3	Liquidity4	DSCR9	Leverage10	CIC7

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

		14.06%	17.44%	-19.53%	6.87%	15.67%	-11.76%	-14.67%
7	50.68%	Return14n 11.48%	Profitability2 19.5%	Efficiency3 -18.62%	Liquidity4 5.11%	DSCR9 20.07%	Leverage20 -10.64%	CIC7 -14.58%
8	50.68%	Return14n 10.83%	Profitability13 18.25%	Efficiency3 -19.94%	Liquidity16 9.64%	DSCR9 18%	Leverage20 -10.04%	CIC7 -13.3%
9	50.67%	Return14n 14.36%	Profitability4 16.33%	Efficiency3 -19.51%	Liquidity4 7.27%	DSCR9 15.82%	Leverage10 -12%	CIC7 -14.72%
10	50.64%	Return14n 13.14%	Profitability2 15.09%	Efficiency3 -18.61%	Liquidity16 7.85%	DSCR9 20.15%	Leverage4 -11.85%	CIC7 -13.3%
11	50.62%	Return14n 12.02%	Profitability4 17.73%	Efficiency8 -18.5%	Liquidity4 5.28%	DSCR9 21.14%	Leverage4 -11.26%	CIC7 -14.06%
12	50.60%	Return14n 11.55%	Profitability13 18.46%	Efficiency3 -19.59%	Liquidity16 8.01%	DSCR9 16.59%	Leverage1 -12.07%	CIC7 -13.74%
13	50.58%	Return14n 11.72%	Profitability4 18.39%	Efficiency3 -18.6%	Liquidity4 5.48%	DSCR9 20.32%	Leverage20 -10.85%	CIC7 -14.63%
14	50.58%	Return14n 11.88%	Profitability13 22.24%	Efficiency8 -19.13%	Liquidity17 7.11%	DSCR10 14.66%	Leverage4 -8.99%	CIC7 -15.99%
15	50.53%	Return14n 13.38%	Profitability4 14.12%	Efficiency3 -18.61%	Liquidity16 8.11%	DSCR9 20.39%	Leverage4 -12.07%	CIC7 -13.33%
16	50.47%	Return14n 11.26%	Profitability13 21.34%	Efficiency8 -20.37%	Liquidity4 8.05%	DSCR9 15.78%	Leverage10 -9.11%	CIC7 -14.09%
17	50.44%	Return7 16.49%	Profitability13 17.16%	Efficiency8 -18.52%	Liquidity4 5.01%	DSCR9 18.57%	Leverage20 -10.65%	CIC7 -13.6%
18	50.44%	Return14n 14.63%	Profitability13 16.74%	Efficiency3 -21.22%	Liquidity16 9.18%	DSCR21 11.25%	Leverage10 -11.21%	CIC7 -15.78%
19	50.42%	Return14n 9.71%	Profitability2 20.55%	Efficiency8 -19.79%	Liquidity4 6.24%	DSCR9 19.93%	Leverage20 -9.76%	CIC7 -14.01%
20	50.42%	Return10 -16.65%	Profitability6 -9.29%	Efficiency2 -14.25%	Liquidity11 14.27%	DSCR2 16.16%	Leverage19 -2.59%	CIC7 -33.54%

The percentage below the name of each risk factor is the weight of that risk factor in the scorecard model.
Highlighted is the model ranked number one in AR, which is also selected as our scorecard model.

Table 15: Top 20 scorecard models for Large Corporations with at least 5% absolute value of weights

Since there are no special drawbacks for the risk factors enter the model with the highest AR for Large Corporations, we suggest using the model ranked number one in AR on Table 15. Note that the weights and AR of our model are not yet fixed. In Session 3.9, we will do the weight adjustment to refine the weight of each category to improve the intuitiveness and practical usefulness of the scorecard. Before that, we will do the cross-validation in Session 3.8 to validate our model selected. If the model passes the cross-validation, we can justify the validity of the risk factors we proposed, and so no need to do further changes in considering other combinations of risk factors.

Worksheet 13 contains all the combinations of risk factors selected during the best-subset selection. The filtering of the weights and sorting in AR are already done in the files.

MFA – Large Corporate.xlsx

Worksheet 13: Multi-factor analysis for Large Corporations

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

3.8. CROSS-VALIDATION

In Statistics, after we build a regression-type model, we need a procedure called model validation to validate the prediction ability of the model. The key rationale of model validation is to test the prediction performance of the model to the new observations out of the sample used to develop the model. In practice, it is common a model performs well in prediction to the development sample, but it performs poorly in application, and this is the reason why model validation makes sense and is necessary.

The basic reason to cause the risk of poor performance in application is that the model can only generalise the pattern appears in the development sample, when we apply the model to some new observations appear outside the “space” covered by the observations in the development sample, or when the development sample itself is lack of representativeness of the population, or when the population is not stable that the pattern in the population changes, these reasons will all cause the model inapplicable to the observations we meet in application. A more serious possibility is the over-fitting problem in Statistics, it means a model actually “memorises” the pattern in the development sample instead of “generalises” the pattern, so that the model fits the development sample too much and is lack of generality to the other observations out of the sample, in this case the prediction ability in application will definitely be poor.

To do the model validation, we can split our sample into two parts, one part called the training sample, it is used to develop the model, another part called the testing sample, used to test the prediction performance of the model; more ideally, the testing sample should be collected independently and separately to the sample on hand (can be the same, overlapping or non-overlapping sampling time period as the original sample, but customers must be different to the original), so that the testing sample will be more likely to cover the different characteristics to our original sample. The above method is called out-of-sample validation. In addition, when the model is used for prediction in the future (like scorecard models), if possible we should also collect the testing sample from a different time period of observation than the time for the training sample (customers can be fully the same, partly the same, or different to the original sample), this is called the out-of-time validation.

However, in practice often to collect out-of-sample or out-of-time data is difficult or impossible (at least up to the time of model development), and even just to set apart a testing sample by splitting the data is not in favor of the model development, because the data remained for the training sample may be not enough when the data on hand is scarce. These are exactly the conditions we face for our scorecard model development, to ensure the model quality, we do not advice the reduce the sample size for fitting the model, as it will reduce the number of “Bad” customers in the sample and reduce the accuracy of the model prediction. To overcome the difficulty of collecting new observations for validation, there is a method developed in Statistics called cross validation, which makes use of all the data on hand, by systematically splitting the data into the training and testing samples, and evaluates the model performance repeatedly, to try to get a picture of the stability of the model prediction performance and the sensitivity of the parameter estimates.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

We will adopt the cross validation procedure to our scorecard model and to justify the stability of the model AR and the weights of the risk factors. The following procedure describes the steps of the cross validation we use:

1. Randomly split the model development sample into two parts, 80% of the observations are allocated to the training sample, and 20% to the testing sample

2. Use the training sample to re-estimate our scorecard model and calculate AR of the model scores from the training sample, and save the weights of the risk factors.

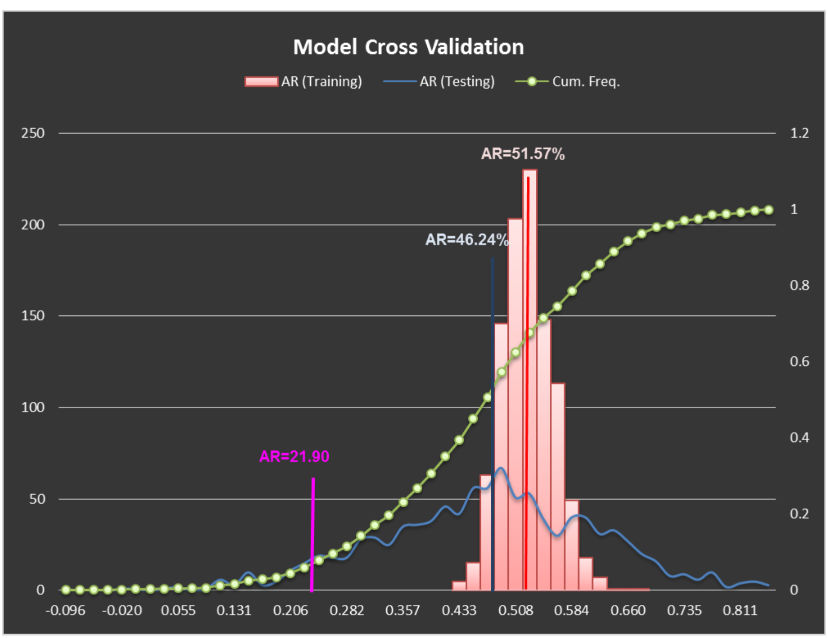
3. Apply the scorecard model fitted to the training sample in step 2 to calculate the scores for the observations in the testing sample, then calculate AR for the scores and Good/Bad flag in the testing sample.

4. Repeat 1000 times from step 1 to step 3, to obtain 1000 different pairs of training-sample AR and testing-sample AR, as well as 1000 different sets of risk factor weights.

5. Plot the distributions of the training-sample AR and testing-sample AR, and the distribution of the weight in absolute value of each risk factor, and calculate the summary statistics (e.g. min, 5-th percentile, mean, 95-th percentile, max, standard deviation).

6. Draw conclusion to the model validity based on the observations in step 5: the model prediction performance is good enough if the shape of testing-sample AR distribution is not concentrated on the left and the 5-th percentile is not too low; the model weights are stable enough if each weight in absolute value is not concentrated to 0 nor concentrated to the right tail, and also the 95-th percentiles are not too large.

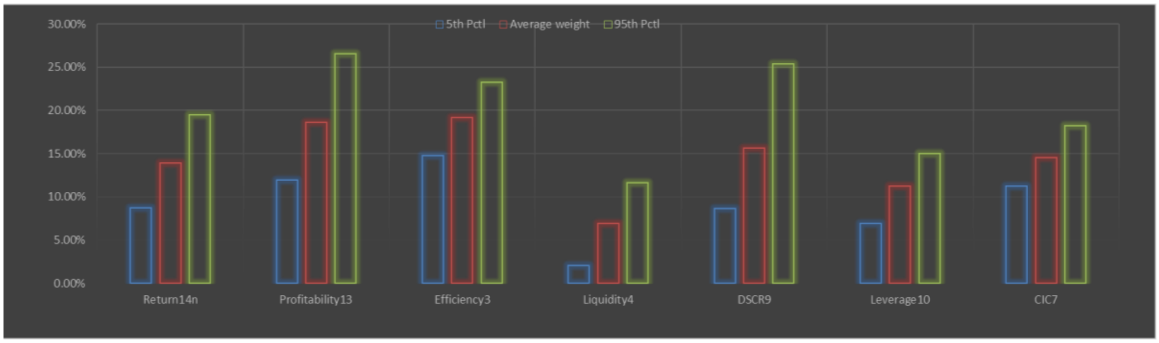
Figure 12 is the plot of training-sample AR and the testing sample AR for Large Corporations; Figure 13 is the plot of risk factor weights in absolute value.



	AR(Training)	AR(Testing)
5th Pctl	46.20%	21.90%
Mean	51.57%	46.24%
95th Pctl	57.30%	69.15%

Copyright © 2014 EntroFine. All Rights Reserved.

Figure 12: Plot of training-sample AR and the testing sample AR for Large Corporations



	Return14n	Profitability13	Efficiency3	Liquidity4	DSCR9	Leverage10	CIC7
Min	1.87%	6.05%	7.77%	0.03%	4.67%	1.90%	7.41%
5th Pctl	8.71%	11.95%	14.80%	2.02%	8.67%	6.90%	11.23%
Mean	13.92%	18.58%	19.20%	6.90%	15.62%	11.23%	14.55%
95th Pctl	19.51%	26.55%	23.26%	11.66%	25.37%	15.04%	18.20%
Max	29.71%	36.79%	27.54%	16.79%	40.13%	17.98%	26.09%
SD	3%	5%	3%	3%	5%	2%	2%

Figure 13: Plot of risk factor weights in absolute value for Large Corporations

After conducting 1000 times of simulation for cross-validation, we conclude that the Large Corporate scorecard model is stable in the prediction performance in terms of AR, and the risk factor weights are stable enough. At the 95% level of confidence, AR of the Large Corporate model for prediction in the worst case still has at least 21.90%. Also at the 95% level of confidence, no weights of risk factors exceed 30% in absolute value. The largest SD of the weights is 5%, the variability is not high.

The following Excel file contains the results of the 1000 times of simulation for the cross validation of Large Corporations.

Cross Validation.xlsx

Worksheet 14: Cross Validation

3.9. WEIGHT ADJUSTMENT

We have justified the validity of the sets of risk factors we choose for our scorecard model. Next, we will do the weight adjustment to it. The rationale and necessity of weight adjustment can be understood at two aspects. First is to restrict the weights of risk factors such that the scorecard do not rely on a particular risk factor by giving too much weight to that risk factor. If the weight on a

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

particular risk factor is too high, that risk factor will dominate the resulting score from the scorecard, other risk factors will not have enough impact to the score. This will cause the generalisation capability of the scorecard being low, and so affect the prediction accuracy of the scorecard in real application. Second is that the weights after adjustment will have better intuitiveness and interpretation, it is because practically we will prefer to round up the weights to integral values in percentage, there will be benefit of better readability of the scorecard model, and hence it can improve the business communication for exchanging the ideas and understanding the realistic operation of the scorecard within the firm.

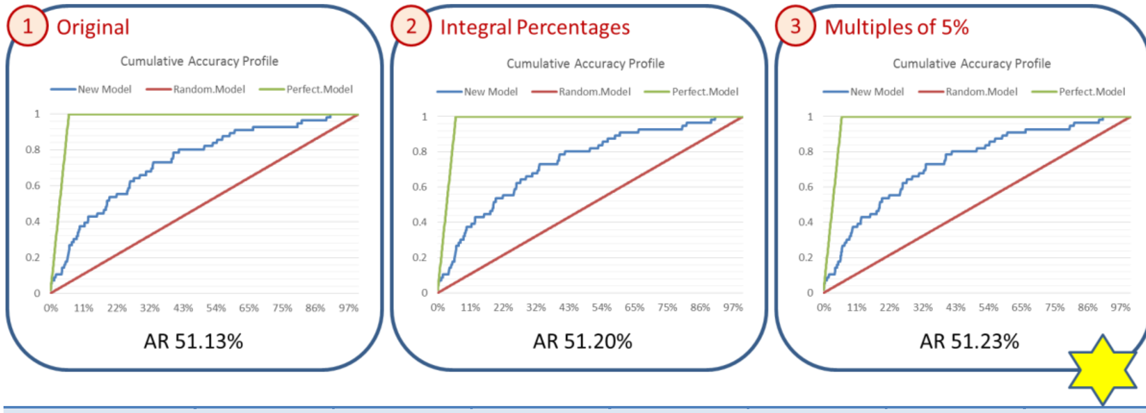
To do the weight adjustment to our scorecard model, we will take two different approaches:

1. Consider all the weights in the model in absolute value. Restrict the weight of CIC7 to 30%, allocate the excess amount of the weight of CIC7 from 30% to the weights of other risk factors proportionally according to the relative proportion of the weights of other risk factors. Then round the new weights to the integral percentage value. Check whether the sum of the new weights equal to 100%, if greater than 100%, then lower the weight which has made the largest amount of increment when rounding up by 1%, and check again if the sum equal to 100%, if still not, then lower the weight which has made the second largest increment when rounding up by 1%, this process is done to the weights which have made the third largest, fourth largest, and so on, if necessary, until the sum equal to 100%; if the sum is lower than 100%, then do the similar process by increasing the weights of those risk factors which have made the largest amount of deduction when rounding down by 1% until the sum equal to 100%. After the above adjustment, move the signs of the original weights to the new weights.

2. Similar to the first approach, consider all the weights in the model in absolute value first. And also restrict the weight of CIC7 to 30% and allocate the excess amount of the weight to the other risk factors by the same approach. Then instead of rounding the new weights to the nearest integral percentage value, this time we round to the nearest multiple of 5 % (i.e. 5%, 10%, 15%, …, 95%). To do so, we first divide the whole interval of [0%, 100%] into the sub-intervals [0%,5%], (5%,10%], …, (95%,100%]. If a weight falls into the first interval, i.e. the weight is lower than or equal to 5%, we will always round up to 5% (we will never round down to 0% as the 0 weight of a risk factor is meaningless). Then for each of the intervals (5%,10%], (10%,15%], …, (95%,100%], we will take the mid-points 12.5%, 17.5%, …, 97.5% as a cutoff point, if the weight is greater than or equal to a cutoff value, it will be rounded up to the upper limit of the interval, similarly if the weight is lower than the cutoff, it will be rounded down to the lower limit. For example, give a weight 27.1%, it belongs to (25%,30%], since it is lower than 27.5%, it will be rounded down to 25%. After the above adjustment, we will check whether the sum of the weights is equal to 100%. If not equal to 100%, we will follow the same logic as the first approach to adjust the weights of those risk factors which have made the largest changes, but the adjustment will be 5% instead of 1%. After ensuring the sum of the weights equal to 100%, we can move the signs of the original weights to the new weights.

Figure 14 plots the CAP curves with the values of AR, and the weights of the risk factors of the original model, and the models with the weights adjusted by the above two approaches. After the weight adjustment, we find that the model adjusted by the second approach (weights are multiples of 5%) performs better than both the model adjusted by the first approach (weights are integral values) and the original model. Hence we propose using the model with weights of multiples of 5% for Large Corporations.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential



	Return14n	Profitability13	Efficiency3	Liquidity4	DSCR9	Leverage10	CIC7
Original	14.05%	18.31%	-19.33%	7.06%	15.09%	-11.57%	-14.59%
Integral Percentages	14.00%	18.00%	-19.00%	7.00%	15.00%	-12.00%	-15.00%
Multiples of 5%	15.00%	20.00%	-20.00%	5.00%	15.00%	-10.00%	-15.00%

Figure 14: CAP curves and weights of the original model and adjusted models for Large Corporations

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

3.10. PD CALIBRATION

By the weight adjustment in the previous session we have finalised our scorecard model. Hence now if we have a customer with the necessary data of the financial statement, we can calculate the score of that customer by the formula $score = \sum_{i=1}^{k} w_i x_i$, where the w_i is the weight of a risk factor x_i (after the transformations described in Session 3.4). This score will then reflect the credit quality of the customer, and by appropriately selecting a cutoff of the score, we can reject the customer if it has the score lower than the cutoff.

Although the score itself can already fulfill the need of accessing a customer's credit quality and hence forms the credit approval system, we will often want a PD (probability of default) estimate from the scorecard by establishing a mapping formula from the score to the PD. In fact, this is the rationale and mechanism how the logistic regression model works – to access the PD of a customer by the risk factors. We can easily find the natural mapping formula from the logistic regression model directly:

$$p = \frac{1}{1+\exp(-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k))}$$

$$= \frac{1}{1+\exp\left(-\beta_0+\sum_{i=1}^{k}|\beta_i|(w_1 x_1+w_2 x_2+\cdots+w_k x_k)\right)}$$

$$= \frac{1}{1+\exp(-\beta_0+\sum_{i=1}^{k}|\beta_i|\times score)}$$

Note that the above formula is only applicable to the scorecard model before weight adjustment. After weight adjustment, w_i is no longer equal to $-\beta_i/\sum_{i=1}^{k}|\beta_i|$ and so the relationship on the above formula breaks down. However, we can easily build up again the relationship between PD and the score by fitting the Good/Bad flag against the score with the logistic regression model:

$$p = \frac{1}{1+\exp(-\alpha-\beta\times score)}$$

By estimating α and β in the above formula, we get the mapping function. It is worth to note that $\alpha = \beta_0$ and $\beta = -\sum_{i=1}^{k}|\beta_i|$ if we fit the Good/Bad flag against the score of the original model without weight adjustment, hence this method is consistent with the natural mapping formula from logistic regression we mentioned above.

The above methods give the natural and direct approaches to build up the relationship between score and PD. However, there is one drawback. Since the sample we use to develop the scorecard model is only some observations from the population, there will be sampling error and it is possible there are some other causes to distort the representativeness of the sample to the population. In this case, our model PD estimates will only be able to reflect the level observed in the sample, there is a risk that they will depart from the PD levels in the whole population. To solve this matter, we will do calibration to the PD of the model. In general, calibration means any methods to adjust the model based on some given real data such that the performance of the model is consistent to the pattern expressed by that real data. To do the PD calibration in our case, we will first estimate the long-run average PD at the population level, then apply the Scaled Logistic Method, it calibrates the PD formula $p = \frac{1}{1+\exp(-\alpha-\beta\times score)}$ by adding a parameter κ which

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

is dependent on the long-run average PD, so the formula becomes $p = \frac{1}{1+\kappa\exp(-\alpha-\beta\times score)}$.

3.10.1.SCALED LOGISTIC METHOD

We will now introduce the Scaled Logistic Method for PD calibration of our scorecard model. To begin with, we first get an estimate of the long-run average PD for Large Corporations. Note that the long-run average PD estimates for SME and MSMEs are 4% and 6% respectively. Since Large Corporations generally have better credit qualities than SME and MSMEs, we select 3% as a referencing long-run average PD estimate for Large Corporations, it should have good enough conservativeness and feasibility.

With this long-run average PD estimate, we can define the parameter κ,

$$\kappa = \frac{1-CT}{CT} \cdot \frac{B}{G}$$

, where CT is the long-run average PD (also called the central tendency), B is the number of “Bad” customers in the sample, G is the number of “Good” customers in the sample

Then by fitting the Good/Bad flag to the scores of the model with weights of multiples of 5% we determined in Session 3.9, we estimate the parameters α and β of the logistic regression model

$$p = \frac{1}{1+\exp(-\alpha-\beta\times score)}.$$

And we can define the following calibrated PD mapping function:

$$p = \frac{1}{1+\kappa\exp(-\alpha-\beta\times score)}$$

Table 16 summarises the parameters we use for the Scaled Logistic Method.

B	G	CT	κ	α	β
56	875	3%	2.0693	-3.2055	-0.0484

Table 16: Parameters for Scaled Logistic Method

Figure 15 plots the calibrated PD against score for Large Corporations.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Calibrated PD vs Score

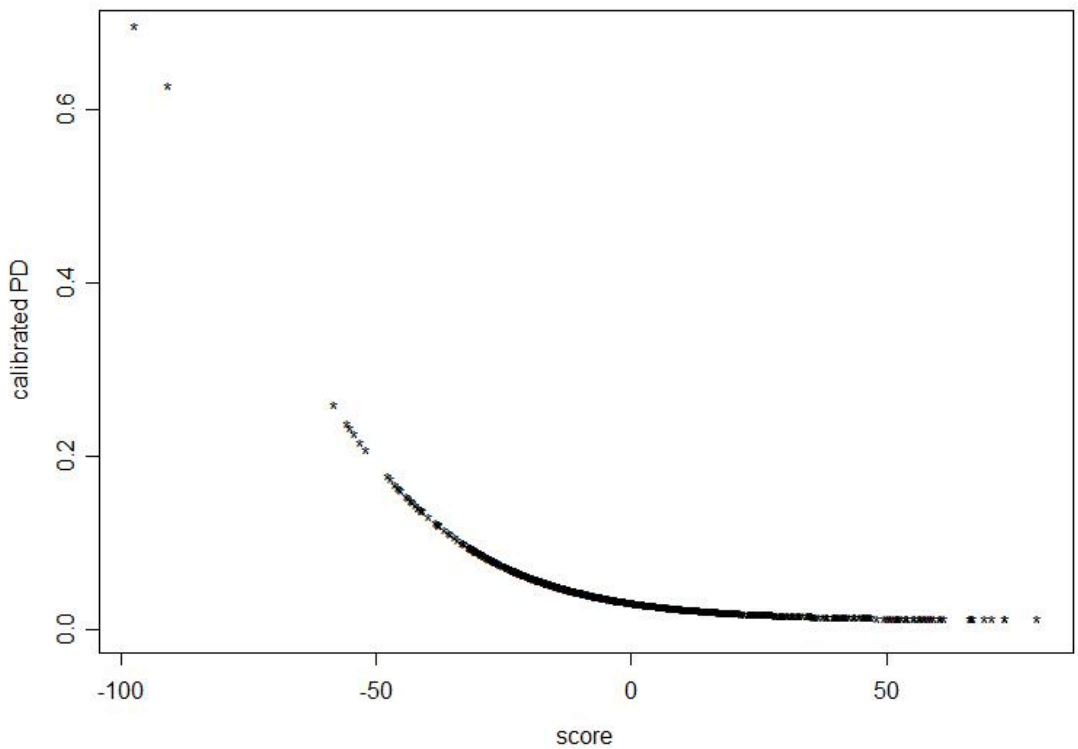


Figure 15: Plot of calibrated PD against score for Large Corporations

3.11. RATING MASTER SCALE

For each customer, in addition to the score and calibrated PD we can obtain from the scorecard, we can also design a master scale which maps a PD to a credit rating according to the PD level, so that each customer will also obtain a credit rating. In industry, usually we will design two different kinds of rating master scale, one is based on the external ratings, the popular choices are the Standard & Poor's (S&P) and Moody's ratings; another one is to construct internal ratings, it means the ratings are used internally in the firm, but usually the mapping between the internal ratings and the external ratings will be established at the same time, so that each internal rating can also reflect the corresponding level on the external ratings scale.

Internal Ratings		S&P Ratings	Moody's Ratings	PD Range (Level 1 Rating Scale)			PD Range (Level 2 Rating Scale)		
Level 1	Level 2			PD-Low	PD-Mid	PD-High	PD-Low	PD-Mid	PD-High
1	1.1	BB+ or above	Ba1 or above	0.00%	0.88%	1.19%	0.00%	0.57%	0.72%
	1.2	BB	Ba2				0.72%	0.90%	1.19%
2	2.1	BB-	Ba3	1.19%	1.96%	2.64%	1.19%	1.57%	1.87%

	2.2						1.87%	2.22%	2.64%
3	**3.1**	B+	B1	2.64%	3.56%	4.66%	2.64%	3.14%	3.59%
	3.2						3.59%	4.09%	4.66%
4	**4.1**	B	B2	4.66%	5.94%	7.70%	4.66%	5.31%	6.02%
	4.2						6.02%	6.81%	7.70%
5	**5.1**	B-	B3	7.70%	9.64%	12.46%	7.70%	8.70%	9.82%
	5.2						9.82%	11.07%	12.46%
6	**6.1**	CCC+	Caa1	12.46%	15.49%	19.35%	12.46%	13.99%	15.62%
	6.2						15.62%	17.41%	19.35%
7	**7.1**	CCC or below	Caa2 or below	19.35%	24.37%	100.00%	19.35%	21.46%	23.76%
	7.2						23.76%	26.23%	100.00%
8	**8.1**	Default Grade		100%			100.0%		
	8.2								
	8.3								

Table 17: Rating master scale

Table 17 shows the rating master scale we propose. The scale consists of both the external ratings (S&P and Moody's) and internal ratings (Level 1 and Level 2), where the Level 2 rating scale is the sub-scale of Level 1, i.e. the Level 1 ratings are from 1 to 8, while the Level 2 ratings add the suffixes .1, .2 to each of 1 to 8 of Level 1 ratings, and also .3 solely to the rating 8. To use the table, we need to read the values of PD-low and PD-High on either the Level 1 rating scale or the Level 2 rating scale, except the last 3 rows of Level 1 rating 8, each row on the table will represent a half-closed interval [PD-low,PD-High), so that when a given PD lies on the interval [PD-low,PD-High), i.e. PD-low ≤ given PD < PD-High, the corresponding Level 1 or Level 2 rating as well as the S&P and Moody's ratings will be assigned to the customer of that given PD.

For the Level 1 rating 8, the sub-ratings 8.1, 8.2 and 8.3 at Level 2 correspond to the groups 3, 4 and 5 respectively of the debt classification published by SBV in Circular 02. Note that the debt classification is the current knock-out factor used by VPBank, where each customer who has debts of the groups 3, 4 or 5 at the time of observation will be rejected. This rule will also be kept in the credit approval system we propose (see Session 4.1), and when the rejection condition is met, we will directly assign the Level 1 rating 8 and the corresponding Level 2 rating 8.1 (group 3), 8.2 (group 4) and 8.3 (group 5) to that customer, and the PD estimate will be taken 100%.

Figure 16 and Figure 17 are the plots of the Level 1 and Level 2 rating distributions for Large Corporations after assigning the ratings to the calibrated PD by our master scale. More than 90% of VPBank Large Corporate customers get the Level 1 ratings 1 – 4. The mode of the Level 1 rating of VPBank Large Corporate customers is 2, which corresponds to BB- of S&P ratings, which is the same to the current country rating of Vietnam rated by S&P.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

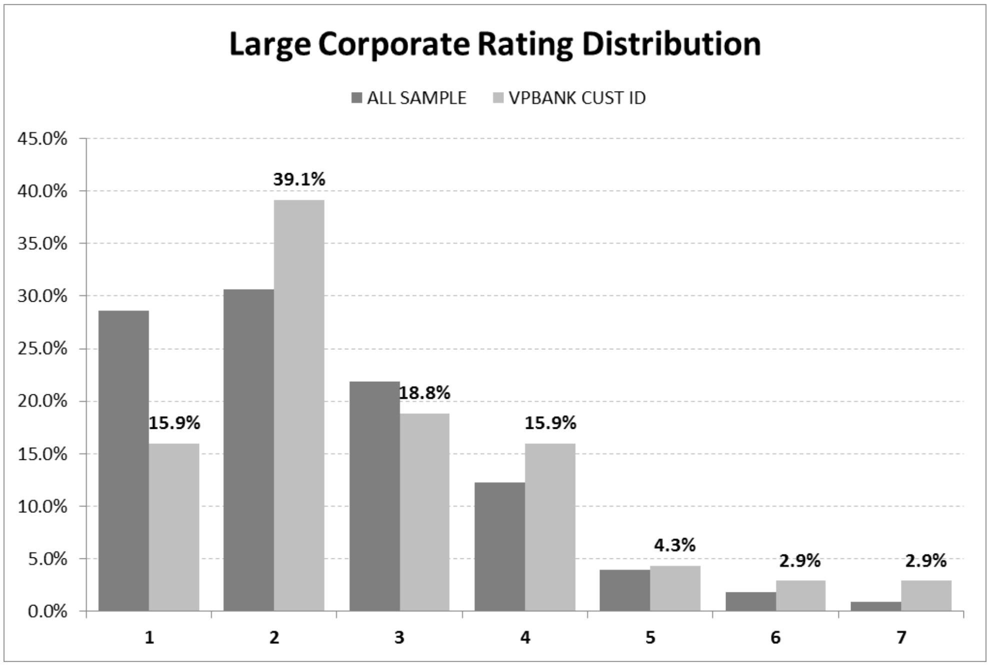


Figure 16: Level 1 Rating Distribution for Large Corporations

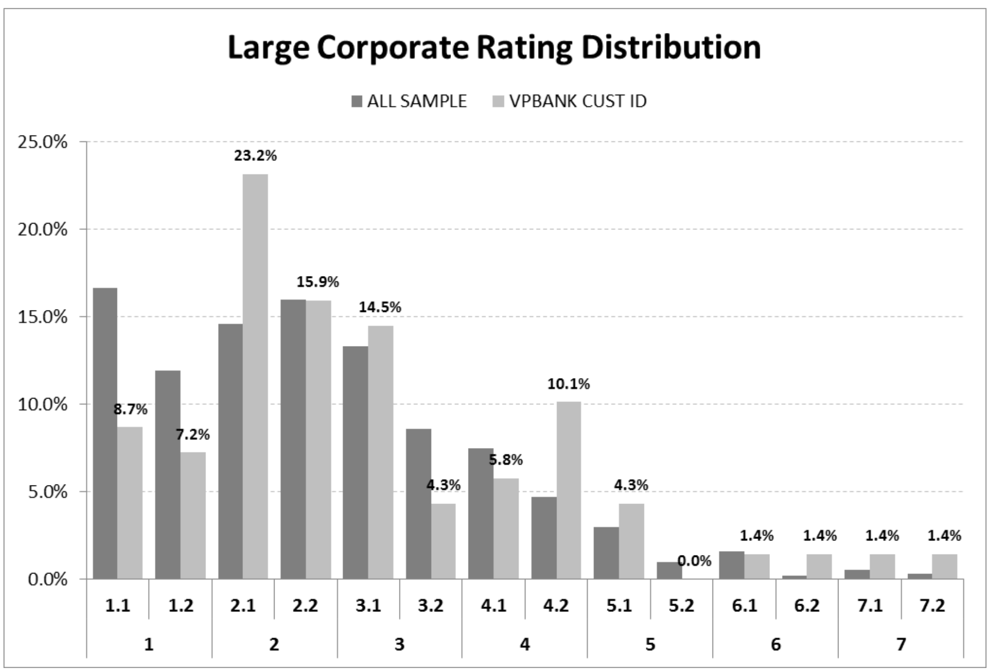


Figure 17: Level 2 Rating Distribution for Large Corporations

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Worksheet 15 shows the score for each customer in the sample, illustrates how to calculate the calibrated PD by the scaled logistic method, as well as how to map the internal and external ratings from the rating master scale.

Score, PD & Rating.xlsx

Worksheet 15: Scores, calibrated PDs and ratings

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

4. MODEL SUMMARY

At the moment we have finished the development of the scorecard, the mapping functions from the score to the calibrated PD, as well as the rating master scale to map from the PD to the internal and external ratings. This session we will summarise our results, to give a macro picture of the whole credit approval system and the whole model structure of the scorecard.

4.1. DECISION TREE OF THE LARGE CORPORATE CREDIT APPROVAL SYSTEM

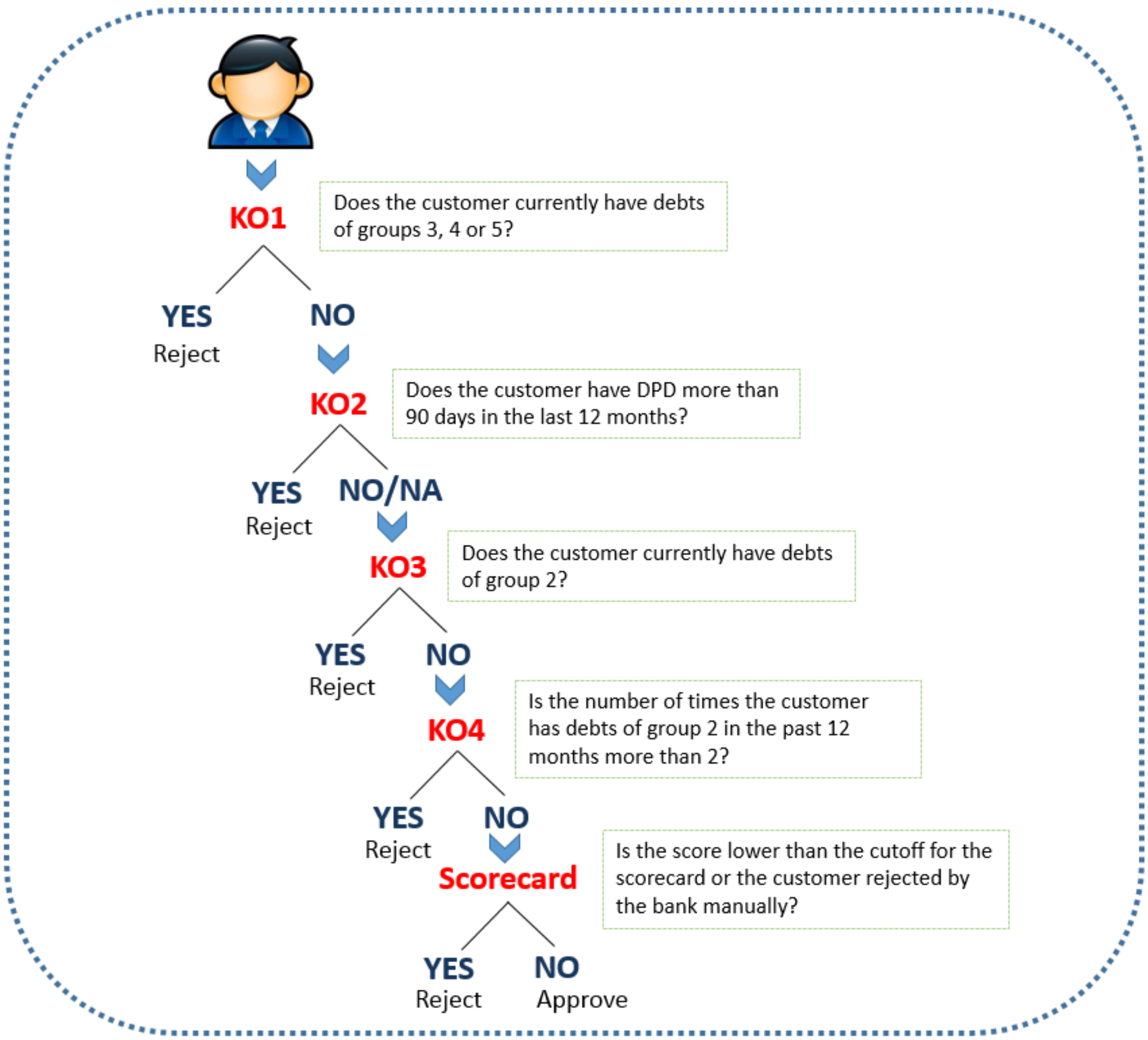


Figure 18: Decision tree of the credit approval system

The credit approval system we have designed consists of two parts, one is the 4 knock-out (KO) factors to directly reject the customers who fulfill the rejection conditions, second is the scorecard

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

which gives each customer a score such that when a customer gets a score below the cutoff score designed for the scorecard, the customer will be rejected.

The first KO factor is the debt classification published by SBV in Circular 02 which we have mentioned in Session 3.11. The other 3 KO factors are CIC9, CIC11 and CIC12, whose rejection rules are described on Table 14. Figure 18 illustrates how the 4 KO factors and the scorecard form the credit approval system, its structure is a decision tree, each KO factor is a decision node, either to reject the customer of not by the decision rule, and lastly the scorecard give the score to the customer for the bank's final approval.

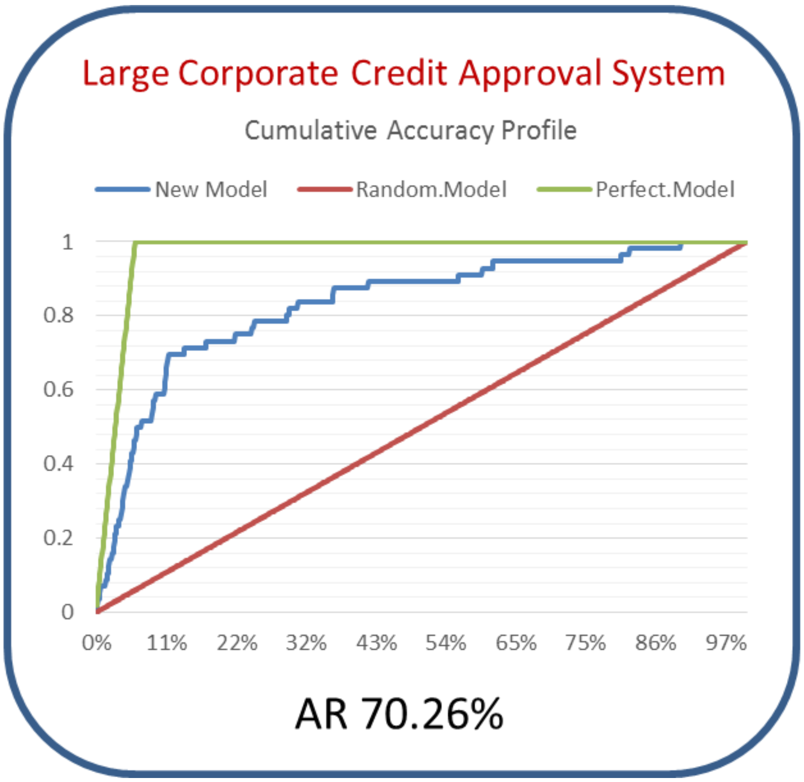


Figure 19: CAP curve of the Large Corporate credit approval system

With the aid of the KO factors, in fact the actual discriminatory power of the whole credit approval system will be much higher than the scorecard solely. To measure how powerful is the whole credit approval system, we can first examine all the observations in the sample by the KO factors (since no observations are of group 3,4 and 5, we will assess only CIC9, CIC11 and CIC12), if a customer fulfills the rejection rule of any KO factor, then the score of that customer will be replaced by the minimum score in the sample, we can then calculate the AR of the new scores, this AR will reflect the discriminatory power of the scorecard together with the KO factors. Figure 19 is the CAP curve and AR of the new scores defined above. We see that the AR of Large Corporations increases to over 70% after the inclusion of KO factors. This evidence is a strong support to the use of the 3 CIC KO factors we have selected on top of the scorecard.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

4.2. MODEL STRUCTURE OF FINAL LARGE CORPORATE SCORECARD

4.2.1. SUMMARY OF FINAL LARGE CORPORATE SCORECARD MODEL

Risk factor	Definition	Weight
Return14n	(Operating Income + Investment Income) / Equity	15%
Profitability13	Profit After Tax / COGS	20%
Efficiency3	Accounts Receivable * 365 / Sales	-20%
Liquidity4	Working Capital / Total Assets	10%
DSCR9	(Cash + Cash Equivalents + Short-term Receivables + Inventories) / (Total Debt + Interest Expenses)	15%
Leverage10	Total Debt / (Total Equity + Long Term Liabilities)	-10%
CIC7	Number of lenders outstanding from the CIC form	-15%

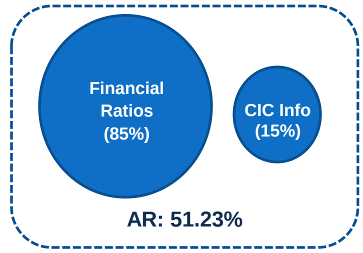


Table 18: Final Large Corporate scorecard model

Table 18 summarises the risk factors we have selected with their definitions and weights in the final scorecard model, we also give the AR and the weight (consider absolute value of weights) of Financial Ratios and CIC Information. So the score of the final scorecard model is calculated by $score = \sum_{i=1}^{k} w_i x_i$, where w_i are the weights, x_i are the risk factors after all the transformations described in Session 3.

4.2.2. EXAMPLES OF HOW TO CALCULATE THE SCORE

Next we will give the examples for the whole calculation process of the scorecard to illustrate the model structure, including the steps from the input the financial/income statement and CIC information, to the transformation process of the risk factors, to the weighted summation to get the final score, as well as the PD calibration and ratting master scale mapping.

Table 19 gives an example of the financial/income statement items and CIC7 which are the necessary inputs for the Large Corporate scorecard. Table 20 illustrates all the intermediate steps of the credit scoring by using the inputs from Table 19, and so calculate the credit score, as well as the calibrated PD and the internal and external ratings.

CODE	FIELD NAME	VALUE
	FINANCIAL STATEMENT	
	Assets	
CT_100	**A. Short-term asset**	651,368,000,000
CT_110	**I. Cash and cash equivalents**	27,592,000,000

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

CT_130	**III. Short-term receivables**	217,273,000,000
CT_140	**IV. Inventory**	405,455,000,000
CT_270	**Total assets**	679,562,000,000
	Equity and Liabilities	
CT_310	**I. Current liabilities**	566,539,000,000
CT_311	1. Short-term borrowings	580,880,000,000
CT_330	**II. Non-current liabilities**	0
CT_334	4. Long-term borrowings	0
CT_400	**B. Equity**	113,023,000,000
	INCOME STATEMENT	
CT_10	**3. Net revenues from goods and services**	1,811,565,000,000
CT_11	4. Cost of sales	1,675,749,000,000
CT_21	**6. Financial revenues**	0
CT_23	- Interest expense	0
CT_30	**11. Net profits from operation**	-527,000,000
CT_60	**19. Profit after corporate tax**	29,175,000,000
	CIC INFORMATION	
CIC7	Q16: What is the number of lenders outstanding from the CIC form?	15

Table 19: Example of financial/income statement and CIC information inputs for Large Corporate scorecard

Step	Description		Return14n	Profitability13	Efficiency3	Liquidity4	DSCR9	Leverage10	CIC7
	Definition		(Operating Income + Investment Income) / Equity	Profit After Tax / COGS	Accounts Receivable * 365 / Sales	Working Capital / Total Assets	(Cash + Cash Equivalents + Short-term Receivables + Inventories) / (Total Debt + Interest Expenses)	Total Debt / (Total Equity + Long Term Liabilities)	What is the number of lenders outstanding from the CIC form?
	Formula		(CT_30+CT_21)/CT_400	CT_60/CT_11	CT_130*365/CT_10	(CT_100-CT_310)/CT_270	(CT_110+CT_130+CT_140)/(CT_311+CT_334+CT_23)	(CT_311+CT_334) / (CT_330+CT_400)	
Start	**Original Value**		0.00	0.02	43.78	0.12	1.12	5.14	15.00
Abnormal Value Treatment / Imputation	**Rule of Abnormal Value Treatment**		when CT_400<=0 then Min	when CT_60<=0 then Max	when CT_10<=0 then Max	when CT_270<=0 then Null	when CT_110+CT_130+CT_140>0 and CT_311+CT_334+CT_23=0 then Max; when CT_110+CT_130+CT_140<=0 and CT_311+CT_334+CT_23=0 then Min	when CT_330+CT_400<=0 then Max	
	Values for Replacing Abnormal Value / Median for Imputation of Missing	**Min**	-962.35				0.02		
		Max		1318.33	16185.45		54824.00	458.23	
		Missing				0.09			
	After Abnormal Value Treatment / Imputation		0.00	0.02	43.78	0.12	1.12	5.14	
Logistic Transformation	**Logistic Transformation Parameters**	**a**	1.78	2.43	3.95	1.46	4.15	3.91	
		b	-5.14	-22.52	-0.03	-9.49	-1.20	-1.46	
	After Logistic Transformation		0.14	0.12	0.08	0.43	0.06	0.97	
Standardisation	**Standardisation Parameters**	**Mean**	0.41	0.24	0.23	0.41	0.25	0.20	7.08
		Stdev	0.25	0.26	0.29	0.28	0.32	0.28	5.10
	Score (After Standardisation)		-54.7722	-24.7979	-25.3989	4.1310	-30.0445	138.8971	77.6473
Credit Scoring	**Weights**		15%	20%	-20%	5%	15%	-10%	-15%
	Weighted Score		-8.2158	-4.9596	5.0798	0.2066	-4.5067	-13.8897	-11.6471
	Total Score		-37.9325						

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

PD Calibration	Scaled Logistic Method Parameters	Alpha	-3.2055
		Beta	-0.0484
		Kappa	2.0693
	Calibrated PD		10.95%
Credit Rating	Level 1 Rating		5
	Level 2 Rating		5.2
	S&P Rating		B-
	Moody's Rating		B3

Light blue: intermediate values of each step of the scorecard; Light orange: outputs of the scorecard

Table 20: Example of the calculation process of the Large Corporate scorecard

We have finished all the contents of the model development and results. Lastly we will introduce the model monitoring and reporting, so that VPBank can follow our methodology to establish its internal policy and process for the monitoring and reporting of the model performance.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

5. MODEL MONITORING & REPORTING

After a scorecard is developed and validated, it will be implemented in a bank's daily business by incorporating the scorecard into the bank's business process (e.g. application scorecards will become the tools of credit approval of new applicants of credit products) with the aid of adequately designed business policies and process flows, as well as information technology (IT).

However, as the time goes on after the scorecard being implemented, the accuracy and performance of the scorecard will have the risk of deterioration, it is on one hand because of the inherent limitation of the methodology of scorecard itself, as any prediction models in science are subject to prediction errors, especially when new observations appear out of the "space" covered by the development sample; on the other hand, the real-world conditions (e.g. economics, political status), are changing every day, and so the structure of the default mechanism in the population is changing, the relevance of the scorecard to the population of observations will decrease as the time goes on.

Because of the above limitation of scorecards, it is feasible and necessary to incorporate a monitoring and reporting system to the business to regularly check the status of the scorecard operation, so that when any problems occur on the scorecard performance, a bank can promptly take remedial actions to tackle the problems, and when necessary the bank may need to consider re-development of a new scorecard.

The model monitoring and reporting approach we propose will cover the **Discriminatory Power** (DP) of the scorecard model (i.e. the ability of the scorecard to separate ex-ante between defaulting and non-defaulting customers), the quality of the model **Calibration** (i.e. the accuracy of the PD estimation) and the model **Stability** in different aspects, including the rating distribution, time stability of discriminatory power, as well as the general conditions for the model use.

5.1. MONITORING & REPORTING OF DISCRIMINATORY POWER

The core function of a scorecard is used for the prediction of the default/non-default status of a customer at a certain time in the future. Hence, the discriminatory power is especially important for the usefulness of a scorecard, and it should be adequately monitored and reported on a regular time basis.

The monitoring of discriminatory power of a scorecard in its nature is similar to the model validation by the use of an out-of-time sample, it is because the sample to be used will be the new observations (may include some customers appeared in the development sample) in the future after the scorecard is implemented in the business, and monitoring itself indeed implies the validation of the model effectiveness for the real operation in business environment.

Accuracy Ratio (AR), which we have used in the model development and validation process as the measure of discriminatory power of a scorecard, will also be used as the key indicator of discriminatory power to be monitored and reported.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

5.1.1. ACCURACY RATIO OF OUT-OF-TIME SAMPLE

To collect the out-of-time sample for monitoring, VPBank have to fix two cutoff dates which are 12 months apart (a common choice in practice is the beginning of a year and the end of a year), so that at the first cutoff date, all the customers on hand under the scope of coverage of a scorecard (i.e. Large Corporations) will be evaluated by the scorecard to get the scores, then 12 months later, their default/non-default statuses reflecting their performance throughout the year will be observed. The following figure illustrates the idea.

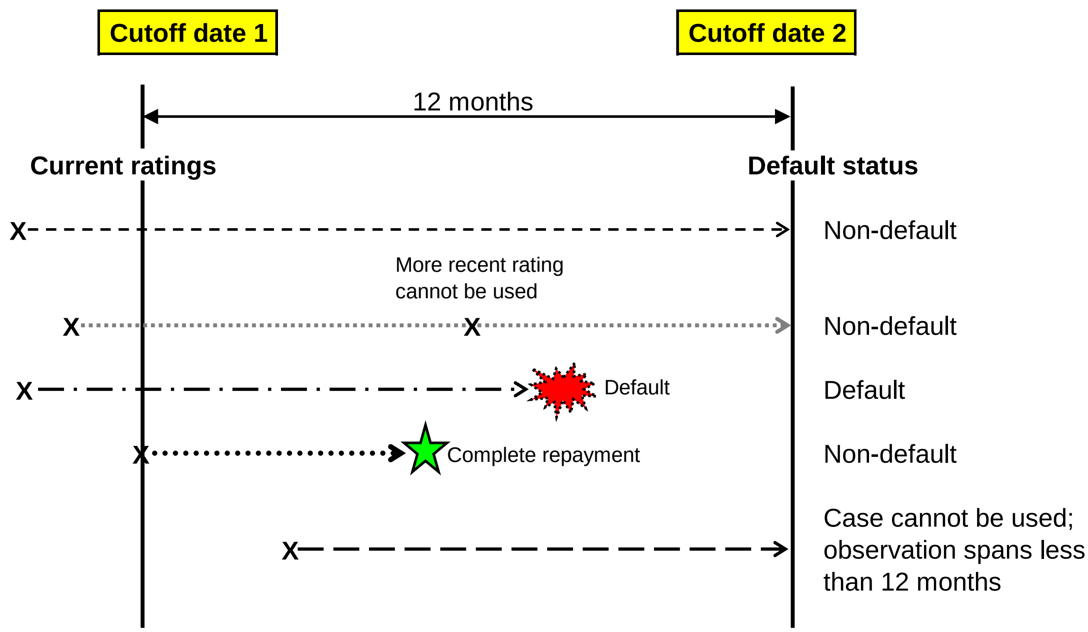


Figure 20: Illustration of collection of out-of-time sample

Then based on the pairs of score and default/non-default status of all customers in the sample for monitoring, AR can be calculated, and also the CAP curve can be plotted for reference. If AR does not reach 30%, then the discriminatory power (AR) of each risk factor should be checked carefully one by one to investigate which risk factors are the source of low discriminatory power. Otherwise if AR is larger than 60%, then it means the scorecard performs well, or if it is between 30% and 60% then it is still acceptable. Figure 21 is a brief process flow chart of the monitoring of discriminatory power of the scorecard.

Note that the above process should be repeated continuously for every year. A clear policy should be documented which requires remedial actions to address the conditions when AR is less than 30%, and any remedial actions should also be documented. After checking the discriminatory power, VPBank can move on to check the validity of PD calibration.

Copyright © 2014 EntroFine. All Rights Reserved.

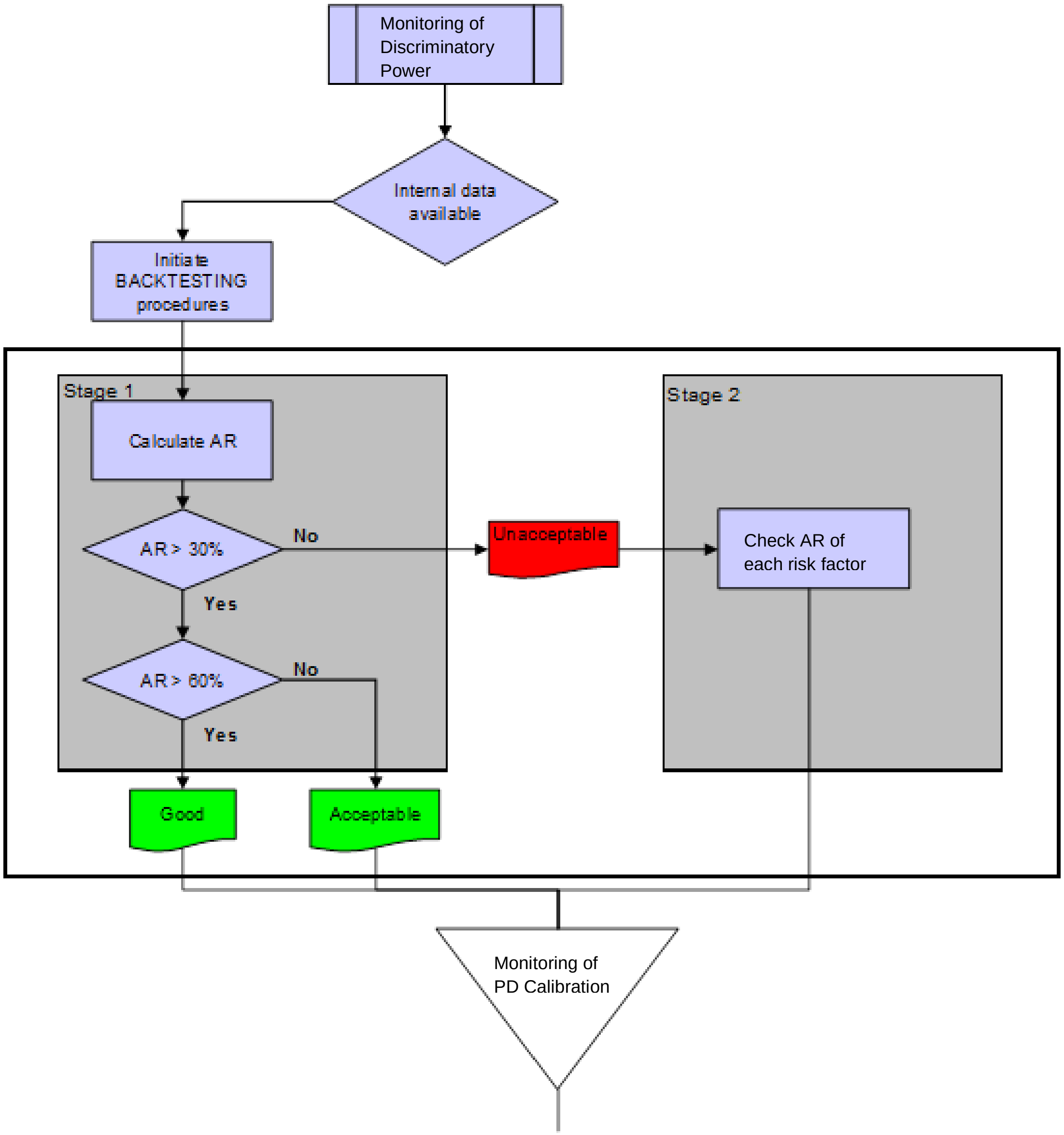


Figure 21: Process flow chart of monitoring of discriminatory power

5.2. MONITORING & REPORTING OF PD CALIBRATION

The second major function of a scorecard is to produce a PD estimate for each customer. So the accuracy of this PD estimate will be another concern of the scorecard effectiveness. To access the accuracy, we can observe how the customers actually perform in reality at the end of a 1-year period from the time of estimation. By measuring the proportion of the actual defaulted customers, which is called the realised default rate, we can use it as a basis to compare with the PD estimate from the scorecard and measure the prediction error, from which we can infer statistically whether

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

the scorecard accurately predicts the real default status by its predicted PD. This statistical procedure to examine the PD prediction accuracy is called PD backtesting. Before going into the details of the methodology of PD backtesting, we need to first introduce some basic theoretical ground regarding the nature of the expected cyclicality of the PD estimates from the scorecards, it will affect the methods we use to carry out PD backtesting.

In theory there are two kinds of expected cyclicality of PD estimates, one is called TTC ("Through-the-cycle" rating philosophy), for which the PD estimates reflect the long-run average level of the customers' default chance throughout the economic cycle, so that there is little reason to expect accuracy of PD estimates in any particular year given that the 1-year realised default rate is directly compared to the PD estimates; another one is called PIT ("Point-in-time" rating philosophy), for which the PD estimates reflect the forecast of the customers' default chance 1 year afterwards, hence a direct comparison of the 1-year realized default rate to the PD estimate is meaningful.

If a scorecard's rating philosophy is TTC, then we will need special adjustment to the 1-year realised default rate before making comparison to the PD estimate in the PD backtesting, the adjustment here is based on statistical theory and parameters for the adjustment will be estimated by real data; if a scorecard's rating philosophy is PIT, then we can directly move to the PD backtesting to use the 1-year realised default rate without any adjustment.

Our Large Corporate scorecard model, as explained in Session 3.10, provides the PD estimates which are calibrated to the long-run average level. Hence they in nature adopt the TTC rating philosophy. Therefore we will adopt the special adjustment to the 1-year realised default rate before moving on to the PD backtesting. Next we will introduce the methodology of doing this special adjustment by mobility metric. To begin with, we will first introduce the rating migration matrix, which is the key input of the parameter estimates determining the degree of adjustment.

5.2.1. RATING MIGRATION MATRIX

Before we explain the rationale of constructing and using the rating migration matrix, we will first give a more detailed explanation for the properties and effects of different rating systems with the different natures of rating philosophy. Figure 22 illustrates the idea.

Theoretically, the PD estimates given by a TTC rating target at the long-run average level, if there are no fundamental changes in the creditworthiness of the obligors, the PD estimates will be stable throughout the economic cycle. In the perfect condition, the PD estimates will be constant as the red line indicates on the figure (Perfect TTC). However, in reality, most TTC ratings will have some degree of fluctuations of the PD estimates, like the Standard & Poor's (S&P) rating, which is recognized as a TTC rating, would have the fluctuations like blue line indicates in the graph. Another famous rating model in the industry, the KMV model, is PIT instead, its PD estimate will vary more significantly from year to year, like the black line indicates in the figure. It is quite common in practice that an internal credit rating model of a bank will lie somewhere between the two extremes of the S&P rating and the KMV model, which respectively represents two standard benchmarks in the industry of the TTC rating and the PIT rating.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

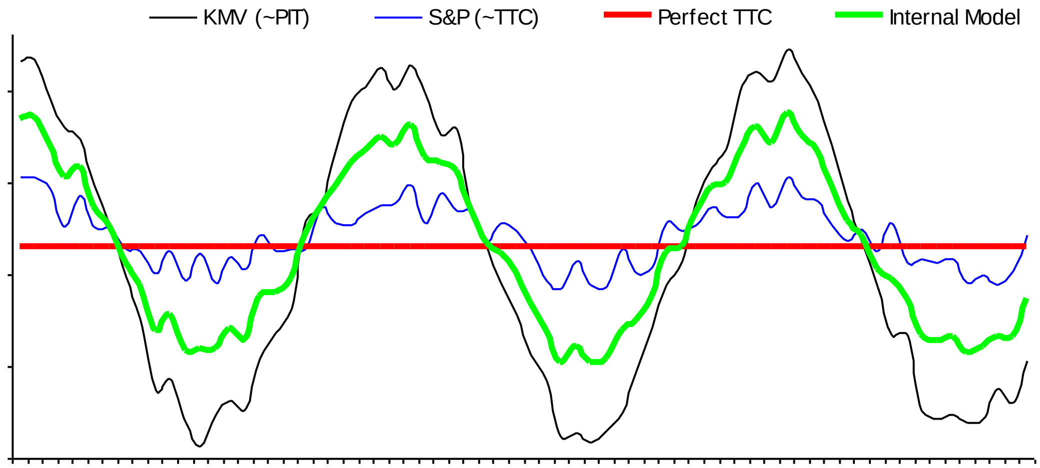


Figure 22: Illustration of the changes in PD estimates by different rating philosophies

Next we will introduce the rating migration matrices of S&P rating and KMV model respectively to facilitate our explanation. Table 21 shows the rating migration matrix for S&P. The legend for the rows mean the original rating of a company, the legend for the columns mean the rating a company changes to 1 year later (D: Default, NR: No rating has been requested, or there is insufficient information on which to base a rating), for each cell of the matrix, each year S&P can measure the proportion of companies transiting from an original rating in the previous year to the new rating of that year, by averaging the transitions rates obtained for each year of the period 1981 to 2005, it will give the number in the cell of the matrix. Note that each row sums into 100 (%), hence each row can be reviewed as the empirical distribution of the different new ratings conditional on each given original rating. Table 22 gives the rating migration matrix for KMV for the period 1990 to 9/1995.

From/To	AAA	AA	A	BBB	BB	B	CCC/C	D	NR
AAA	88.20	7.67	0.49	0.09	0.06	0.00	0.00	0.00	3.49
AA	0.58	87.16	7.63	0.58	0.06	0.11	0.02	0.01	3.85
A	0.05	1.90	87.24	5.59	0.42	0.15	0.03	0.04	4.58
BBB	0.02	0.16	3.85	84.13	4.27	0.76	0.17	0.27	6.37
BB	0.03	0.04	0.25	5.26	75.74	7.36	0.90	1.12	9.29
B	0.00	0.05	0.19	0.31	5.52	72.67	4.21	5.38	11.67
CCC/C	0.00	0.00	0.28	0.41	1.24	10.92	47.06	27.02	13.06

Table 21: S&P global average 1-year transition rates (%), 1981 to 2005 [1]

[1] Source: S&P's report; Annual 2005 Global Corporate Default Study And Rating Transitions

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Initial Rating	Rating at year-end (%)							
	AAA	*AA*	*A*	*BBB*	*BB*	*B*	*CCC*	*Default*
AAA	66.26	22.22	7.37	2.45	0.86	0.67	0.14	0.02
AA	21.66	43.04	25.83	6.56	1.99	0.68	0.20	0.04
A	2.76	20.34	44.19	22.94	7.42	1.97	0.28	0.10
BBB	0.30	2.80	22.63	42.54	23.52	6.95	1.00	0.26
BB	0.08	0.24	3.69	22.93	44.41	24.53	3.41	0.71
B	0.01	0.05	0.39	3.48	20.47	53.00	20.58	2.01
CCC	0.00	0.01	0.09	0.26	1.79	17.77	69.94	10.13

Table 22: KMV model's 1-year transition rates (%), 1990 to 9/1995 [2]

Note that the diagonal elements of the S&P rating migration matrix is much closer to 1 than that of KMV model, it indicates that the S&P ratings are more stable from year to year than the KMV model, this fact is consistent to what the Figure 22 expresses for the TTC nature of the S&P rating and the PIT nature of the KMV model, as the same rating indeed reflects the same level of PD estimation. Hence, by the similar idea, for a particular internal credit rating system of a bank, we can try to construct a rating migration matrix for that internal rating, and then the degree of how the diagonal elements of this matrix close to 1 will reflect the "through-the-cycle-ness" of this internal rating. With some special numerical technique (i.e. the mobility metric, which is calculated from the input of the rating migration matrix) which depends mainly on the diagonal elements, we can measure how stable are the ratings of this internal rating system compared to S&P and KMV respectively. If this internal rating system's stability in rating migration is "closer" to S&P, we will treat its rating nature as more TTC-like, if it is "closer" to KMV, we will treat its rating is more PIC-like in nature.

In fact the above idea is the rationale behind our methodology to make the adjustment to the 1-year realised default rate. As we mentioned previously, typically in practice a scorecard of a bank will lie somewhere between S&P and KMV in terms of its "TTC-ness", so that the adjustment we propose is based on how "close" the scorecard to S&P and KMV in its "TTC-ness". It means, when we find that our Large Corporate scorecard performs more TTC-like as its rating migration stability is "closer" to S&P, more adjustment will be made, when it performs less TTC-like as it is "closer" to KMV, less adjustment will be made. And this idea of adjustment is consistent to our discussion at the beginning of this sub-session regarding the need of adjustment when the rating philosophy is TTC.

To realise the adjustment described above, VPBank has to first construct its own rating migration matrix similar to what S&P and KMV did. VPBank should base on the ratings described on the rating master scale (Table 17) we proposed, for each of Level 1 and Level 2 of internal ratings, as

[2] Source: S&P's report; Annual 2005 Global Corporate Default Study And Rating Transitions

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

well as S&P and Moody's ratings, such rating migration matrix can be constructed respectively, just accumulating the rating measures for the same group of customers for at least 2 years can form the matrix, for the ratings more than 2 years, average yearly migration rates can be used.

Next sub-session we will explain in details how we do the adjustment described above by using the mobility metric with the input of the rating migration matrix.

5.2.2. REALIZED DEFAULT RATE WITH ADJUSTMENT BY MOBILITY METRIC

We will begin with the derivation of the Mobility Metric[3] (MM), which mainly measures how close the diagonal elements of the rating migration matrix are to 1, so can effectively reflect how strong the "TTC-ness" of a credit rating system is.

Suppose the rating migration matrix of a given credit rating system is P, where P should be a square matrix (i.e. n $\times$ n matrix) showing completely all the migration between different ratings, including the "Default" rating on the legend of the rows. Note that the "Default" rating is called the absorbing stage in the Markov chain theory, which once a customer gets a "Default" rating, will not change back to the other ratings (in reality it is possible but very rare, usually in practice we just assume it is impossible). In mathematical words, it means the transition rates from the "Default" rating to the other ratings will be 0, and the transition rate to the "Default" rating itself will be 1. Hence if we put the "Default" rating on the last row of the rating migration matrix, the last row will appear as a row vector with all the leading 0 elements on the left but only an element of 1 in the last cell.

With the rating migration matrix P, we can define the Mobility Matrix $\tilde{P}$ as follows:

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

,where I is the identity matrix with the conforming dimensions

The purpose of subtracting the identity matrix from P above is to make the diagonal of $\tilde{P}$ reflects the difference between the diagonal elements and 1

Then we can define the Mobility Metric (MM) as follows:

$$MM = \frac{\sum_{i=1}^{N} \sqrt{\lambda_i(\tilde{P}'\tilde{P})}}{N}$$

,where $\sqrt{\lambda_i(\tilde{P}'\tilde{P})}$ is called the i-th singular value of $\tilde{P}$, $\lambda_i(\tilde{P}'\tilde{P})$ denotes the i-th eigenvalue of the matrix $\tilde{P}'\tilde{P}$

[3] Jafry & Schuermann, Measurement, estimation and comparison of credit migration matrices, *Journal of Banking and Finance*, 2004

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

Hence MM in fact is defined as the average of the singular values of $\tilde{P}$. Since in mathematics we know that $\sum_{i=1}^{N}\sqrt{\lambda_i(\tilde{P}'\tilde{P})} = Trace\left[(\tilde{P}'\tilde{P})^{1/2}\right]$, where Trace means the sum of the diagonal elements. So it fact MM is the sum of the square root of the diagonal elements of $\tilde{P}'\tilde{P}$, and this value will be smaller when the diagonal elements of $\tilde{P}$ are closer to 0, i.e. the diagonal elements of P are closer to 1. This fact explains why MM is a valid measure of the degree how close are the diagonal elements of the rating migration matrix are to 1, and so the "TTC-ness" of the credit rating system, or more exactly, the stability of the ratings (and so PD estimates) generated by the credit rating system.

By applying the formula of MM to the transition matrices of Table 21 and Table 22 (after completing to square matrices), it yields 0.229 and 0.484 for S&P and KMV respectively. Based on this observation, and the fact in practice that most credit rating systems lie somewhere between S&P and KMV in terms of the "TTC-ness", we suggest the following formula to be used to adjust the 1-year realised default rate:

$$DR_{Adjusted} = wPD_{CT} + (1-w)DR_{Realised}$$

,where $DR_{Adjusted}$ is the adjusted 1-year realised default rate, PD_{CT} is the central tendency (i.e. long-run average PD) implied by the scorecard, $DR_{Realised}$ is the unadjusted 1-year realised default rate, w is the adjustment factor which is determined by MM using the following formula:

$$w = \frac{0.484-MM}{0.484-0.229}$$

From the above formula, we can see $DR_{Adjusted}$ is actually a convex combination of PD_{CT} and $DR_{Realised}$, which depends on the distance of MM to the two points 0.229 and 0.484. When MM is closer to 0.229 (so that the crediting rating system is "closer" to S&P), w will be closer to 1, and the adjustment will be stronger so that $DR_{Realised}$ will be closer to PD_{CT}; when MM is closer to 0.484, w will be closer to 0 (so "closer" to KMV), and the adjustment will be weaker so that $DR_{Realised}$ will be closer to $DR_{Realised}$.

For our Large Corporate scorecard, we can take two different angles to examine the 1-year realised default rate and the central tendency and do the above calculation for the adjustment. First, we can make the calculation on the whole sample on hand, so that PD_{CT} will be the central tendency of our Large Corporate scorecard (i.e. 0.3), $DR_{Realised}$ will be the 1-year realised default rate on the whole sample. Second, we can first choose either one scale of internal or external ratings on the rating master scale we proposed (i.e. Level 1, Level 2, S&P or Moody's), then for each rating class on the scale, we can calculate the 1-year realised default rate, and PD_{CT} will be the corresponding value of PD-Mid of that rating class on Table 17. For the choice of rating migration matrix to calculate MM, we suggest using the matrix consistent to the scale of internal or external rating being examined. When the whole sample is used, then any choice of the rating migration matrix can fulfill the execution of calculation, however, we may choose a higher value of MM to increase the conservativeness of the monitoring. For both the two angles above, in fact we can also determine the value of w judgmentally by the preference of VPBank and its conservativeness to the monitoring of PD calibration, a lower value of w should be used if more

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

conservative so that less adjustment will be made and it will be easier to alarm on the PD backtesting.

For the reporting purpose, VPBank should report both the actual 1-year realised default rate and the adjusted realised default rate, with the predicted PD (PD-Mid on the rating master scale for each rating class; or central tendency for whole sample), for the ratings on a chosen internal or external scale on the rating master scale – Table 17, and in addition for the whole sample.

After we have got the adjusted 1-year realised default rates by the method above, we can apply them as the input to do the PD backtesting, which is a procedure to test the statistical significance of the prediction error of the scorecard PD estimate by testing the rigorously established statistical hypotheses, so that it can alarm the risk of poor PD calibration quality according to the conservativeness of VPBank by controlling the level of significance (which is equivalent to the Type-I error) for the statistical tests.

5.2.3. PD BACKTESTING

No matter for the purpose of sound risk management within a bank or for the supervisors' requirement on the adequate preparation of the regulatory capital, the major concern regarding the PD estimates is that these estimates must not be too low in value. If the PD is underestimated, it directly means the bank underestimates the risk involved in its credit portfolio, and so may take insufficiently strong actions to manage the risk; on the other hand, the regulatory capital calculated will be smaller in amount which is indeed not enough to cover the potential loss, and this is not accepted by the supervisors.

Therefore, when we do the PD backtesting, our focus will lie on testing whether the PD is underestimated or not, this can be achieved by using carefully designed hypothesis testing methods in Statistics. There are 3 methods we propose to use for the PD backtesting:

1. Binomial test with assumption of independent default events
2. Moment matching with assumption of non-zero default correlation
3. Granularity adjustment

For all these 3 methods, the adjusted 1-year realised default rate is used as a test statistic to test the null hypothesis that the real value of the population PD is equal to the PD estimate from the scorecard. The binomial test is the least conservative test among these 3 methods, it is because the independent default events assumption is strong, when this assumption is violated, a relatively weak evidence will be easier to reject the null hypothesis; while moment matching and granularity adjustment relax this assumption by taking asset values' correlation into account (which drives the default events' correlation), and so stronger evidence will be required to reject the null hypothesis than the binomial test. So if the binomial test indicates the PD is not underestimated, the remaining 2 tests need not be conducted as their conclusions probably will be the same. However, if the binominal test indicates that the PD is underestimated, we may conduct the remaining 2 tests and check if the tests give the same conclusion. The conclusion from the latter 2 methods will be more reliable given that these methods overcome the limitation

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

of the binomial test. Next, we will go through all these 3 methods one-by-one with the detailed explanation with mathematical formulas.

1. Binomial test with assumption of independent default events

Suppose p_0 is the PD estimate by the scorecard (i.e. PD_{CT}), $\hat{p}$ is the adjusted 1-year realised default rate (i.e. $DR_{Adjusted}$), p is the true parameter of the 1-year PD in population.

Then we can construct the following hypotheses:

$H_0 : p = p_0$
$H_1' : p > p_0$ (or $H_1'' : p \neq p_0$)

, where H_0 is the null hypothesis indicating the true PD is equal to the PD estimate, H_1' is the alternative hypothesis for an one-tailed test indicating the PD estimate underestimates the true PD, H_1'' is the two-tailed test for reference, indicating the true PD is not at the same level as the PD estimate

To test the hypothesis, we propose using the following test statistic (which is derived from the Central Limit Theorem or the Wald test):

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

, where n is the sample size of the rating class (or whole sample) being examined

For the alternative hypothesis for the one-tailed test, H_1', we reject H_0 if

$$1 - \Phi(z) < \alpha$$

, where Φ is the cumulative distribution function (c.d.f.) of the standard normal distribution, α is the level of significance pre-specified by VPBank (the value is not necessarily chosen at the general rule of thumb 0.05, it can be determined by the preference and conservativeness of VPBank)

As a reference, if instead the alternative hypothesis for the two-tailed test, H_1'', is used, then we reject H_0 if

$$1 - \Phi(|z|) < \alpha/2$$

Note that in theory actually z proposed above does not really follow standard normal distribution, but it is compared to the standard normal distribution so as to reflect the effect of the adjustment to the realised default rate. To see the effect of the adjustment, we can consider
$\hat{p} = wp_0 + (1 - w)DR_{Realised}$

Hence, z can be expressed as

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

$$z = \frac{wp_0+(1-w)DR_{Realised}-p_0}{\sqrt{p_0(1-p_0)/n}} = (1-w)\frac{DR_{Realised}-p_0}{\sqrt{p_0(1-p_0)/n}}$$

, where the term $\frac{DR_{Realised}-p_0}{\sqrt{p_0(1-p_0)/n}}$ is in fact standard normally distributed under H_0, the factor $1-w$ gives a "penalty" to the term so as to "adjust down" its absolute value, when we compare the whole term, z, to standard normal distribution, the adjustment will take effect so that it will be less easy to reject H_0, and this is indeed the rationale of our adjustment and the effect we want to achieve for a more conservative PD backtesting

2. Moment matching with assumption of non-zero default correlation

In moment matching method, we relax the assumption of independency of the binomial test by introducing a constant asset correlation, i.e. the correlations of the asset values of any two customers are the same (and so the default events' correlations will be the same). In addition, we assume the default rate follows a beta distribution. From the derivation of Tasche[4] 2003, the following shows how we construct the PD backtesting based on the moment matching approach:

i. Calculate $\mathrm{E}[R_n]$ and $\mathrm{var}[R_n]$,

$$\mathrm{E}[R_n] = p_0$$
$$\mathrm{var}[R_n] = \frac{n-1}{n}\Phi_2(t,t,\rho) + \frac{p_0}{n} - p_0^2$$

,where

$R_n =$ 1-year realised default rate (in the form of random variable)

$p_0 =$ PD estimate (of a rating class or whole sample)

$n =$ number of obligors (of a rating class or whole sample)

$t = \Phi^{-1}(p_0)$

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

$\Phi =$ standard normal cumulative distribution function (c.d.f.)

$\rho =$ asset correlation

On the above formula, for the estimation of the asset correlation, we may use the estimate from the Basel II capital formula[5] as a reference value. The formula is

$$\rho = 0.12 \times \frac{1-e^{-50p_0}}{1-e^{-50}} + 0.24 \times \left(1 - \frac{1-e^{-50p_0}}{1-e^{-50}}\right)$$

[4] Dirk Tasche, A traffic lights approach to PD validation, 2003
[5] P. 64, Internal Coverage of Capital Measurement and Capital Standards, Bank for International Settlements, June 2006

The value of this ρ ranges from 12% to 24%, depending on the PD estimate, i.e. the higher the PD estimate, the lower is the asset correlation.

ii. Calculate the parameters for beta distribution:

$$a = \frac{\mathrm{E}[R_n]}{\mathrm{var}[R_n]}\left(\mathrm{E}[R_n](1-\mathrm{E}[R_n]) - \mathrm{var}[R_n]\right)$$

$$b = \frac{1-\mathrm{E}[R_n]}{\mathrm{var}[R_n]}\left(\mathrm{E}[R_n](1-\mathrm{E}[R_n]) - \mathrm{var}[R_n]\right)$$

iii. Test the hypothesis:

Suppose $\hat{p}$ is the adjusted 1-year realised default rate, we can test the hypotheses H_0 and H_1' as described in the binomial test by using the follow rejection rule:
reject H_0 if $1 - F_{beta}(\hat{p}; a, b) < \alpha$

, where F_{beta} denotes the c.d.f. of beta distribution, a, b are the parameters calculated from ii, α is the level of significance chosen

As a reference if H_1'' is tested instead of H_1', the rejection condition to H_0 will become $\min\left(F_{beta}(\hat{p}; a, b), 1 - F_{beta}(\hat{p}; a, b)\right) < \alpha/2$

Similar to binomial testing, the adjustment by MM will make H_0 less easy to be rejected than using the unadjusted realised default rate directly.

3. Granularity adjustment

Granularity adjustment is an alternative to moment matching, which also takes into account the asset correlation. The key difference to moment matching is that the assumption for the distribution of the realised default rate is normal distribution instead of beta distribution.

To test the hypotheses H_0 and H_1' , we use the following rejection rule:

reject H_0 if $\hat{p} > Q + \frac{1}{2n}\left(2Q - 1 - \frac{Q(1-Q)}{\Phi\left(\sqrt{\frac{\rho\Phi^{-1}(1-\alpha)-t}{\sqrt{1-\rho}}}\right)}\left(\frac{(1-2\rho)\Phi^{-1}(1-\alpha)-t\sqrt{\rho}}{\sqrt{\rho\,(1-\rho)}}\right)\right)$

, where

$\hat{p}$ = adjusted 1-year realised default rate

n = number of obligors (of a rating class or whole sample)

$t = \Phi^{-1}(p_0)$, where p_0 = PD estimate (of a rating class or whole sample)

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

$\Phi =$ standard normal cumulative distribution function (c.d.f.)

$\rho =$ asset correlation

$\alpha =$ level of significance

The above 3 methods provide a good enough framework for VPBank to do the PD backtesting. VPBank should establish its clearly documented policy regarding the monitoring the reporting of the PD calibration accuracy described in this session, including the parameters it uses for the analysis, the contents and objectives to be reported, as well as the remedial actions when conclusion of underestimation of PD is drawn from PD backtesting, etc.

To give an idea to VPBank how is the industry standard for the good practice of PD calibration monitoring and reporting, next sub-session we will describe the expectation of the Hong Kong Monetary Authority (HKMA) to the banks for the validation of PD calibration.

5.2.4. HKMA EXPECTATION TO VALIDATION OF PD CALIBRATION

The following HKMA's expectation is for VPBank's reference only, VPBank can adjust its own expectation according to its capability and the practical conditions. Moreover, VPBank should also keep in mind if SBV issues different expectation, then VPBank has to adjust accordingly.

The HKMA expects banks to establish internal tolerance limits for the differences between the forecast PD and the realised default rates. Banks should have a clearly documented policy that requires remedial actions to be taken when policy tolerances are exceeded, and any remedial actions should also be documented.

Banks should construct the tolerance limits (and the associated policy on remedial actions) around the significance levels used in the statistical tests.

The HKMA expects banks to demonstrate that the internal tolerance limits and remedial actions are commensurate with the risk that the computed capital requirement would not be adequate to cover the default risk incurred. In setting its internal standards, and determining any remedial actions, a bank should be able to demonstrate that it has taken into account a range of factors, including, but not limited to, the relative sizes of the portfolios to which the internal rating systems are applied, the bank's risk appetite in respect of the portfolios, the distribution of the portfolios amongst rating grades, and the inherent risk characteristics of the portfolios.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

5.3. MONITORING & REPORTING OF MODEL STABILITY

The last core performance of the scorecard model which requires regular monitoring and reporting is the stability of the model, where the stability can be interpreted at different aspects, but commonly means the robustness of the model performance with respect to the changes of the factors or conditions underlying the default mechanism in the population.

Actually the discriminatory power and PD calibration accuracy of the model as time moves on can also be viewed as some requirements to the model stability. Because of their importance we separately explained for these two aspects in the previous sessions. There are 3 more aspects of stability that are worth special concern of a bank, they are respectively

1. Stability of the rating distribution
2. Time stability of the discriminatory power, i.e. stability in discriminatory power given forecasting horizons of varying length or as loans become older
3. Stability to changes in the general conditions underlying the use of the model and their effects on individual model parameters and on the results the model generates

We will explain the above 3 types of stability one-by-one in the following.

5.3.1. STABILITY OF RATING DISTRIBUTION

In Session 5.2.1 we explained how to record the migration of ratings by the rating migration matrix, and in Session 5.2.2 we explained how to measure the “TTC-ness” of the rating system by mobility metric. In fact, the “TTC-ness” is a kind of stability of the rating, but it refers to the stability of the rating of a single customer, or to the same group of customers. In a more macro view, we can observe the stability of the whole rating distributions from the sample on hand each year, which includes the old customers in the last year as well as the new customers of the current year, and from the sample we can get a reference to the population stability.

Hence, in addition to the rating migration matrix, VPBank should also monitor and report the rating distribution of the customers on hand regularly. By doing so VPBank can observe whether there are fundamental changes to the creditworthiness of its customer base, and when necessary remedial actions can be taken to control the risk.

5.3.2. TIME STABILITY OF DISCRIMINATORY POWER

In Section 5.1.1, we described the monitoring and reporting of the discriminatory power of the scorecard model over a time horizon of 12 months. However, it is also possible to measure the discriminatory power over longer periods of time if a sufficient dataset is available. In this context, AR will be kept to be used as the key measure of discriminatory power.

As our scorecard model is optimized for a period of 12 months at the development stage, its discriminatory power probably would decrease for longer time horizons. Here it is necessary to ensure that the discriminatory power of the scorecard deteriorates steadily only, that is, without dropping abruptly to excessively low values. Sound scorecard models should also demonstrate

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

sufficient discriminatory power over forecasting horizons of three or more years. To measure the discriminatory power of the scorecard for longer forecasting horizons, the same methodology described in Session 5.1.1 and as illustrated by Figure 20 follows, only except that 12 months is replaced by the longer time horizon decided.

Another aspect of the concern on the time stability of scorecard models is the decrease in discriminatory power as loans become older. This is especially relevant in the case of application scores where it is quite often the discriminatory power for an observed quantity of new business cases decreases noticeably over a period of 6 to 36 months after an application is submitted. This is due to the fact that the data used in application scoring become less relevant over time. Therefore, practitioners frequently complement application scoring models with behavior scoring models. The latter models evaluate more recent information from the development of the credit transaction and therefore provide a better indicator of creditworthiness than application scoring models alone. However, behavior scoring is not possible until a credit facility has reached a certain level of maturity, that is, once behavior-related data are actually available. To access the time stability of discriminatory power as loans getting old, we can compare the ARs of different time horizons for the same group of customers.

5.3.3. STABILITY OF GENERAL CONDITIONS FOR MODEL USE

The assessment of changes in the general conditions under which a scorecard model is used has strong qualitative elements. On one hand, it is necessary to review whether developments in the economic, political, or legal environment will have an influence on the scorecard models or individual model parameters and criteria. On the other hand, internal factors at a bank such as changes in business strategies, the expansion of activities in certain market segments, or changes in organizational structures may also affect the performance of a scorecard model substantially.

Changes in the economic environment include the business cycle in particular, which can cause major fluctuations in the PD parameter during periods of recovery and decline. However, factors such as technical progress or political and legal developments can also influence the effectiveness of the scorecard models.

In particular, country and regional ratings depend heavily on changes in the political environment. Examples in the legal environment include changes in commercial law or accounting standards which may have a positive or negative influence on the effectiveness and significance of certain financial indicators.

Quantifying the effects of changes in general conditions on the functionality of scorecard models requires an in-depth analysis of the model parameters and should therefore accompany the ongoing development of the model. Scorecard models have to undergo further development whenever their performance decreases due to changes in general conditions. On the other hand, VPBank may also decide to develop a new scorecard model if experts believe that a potential or planned change in general conditions would lead to a substantial loss in the performance of the current model.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

5.3.4. HKMA EXPECTATION TO MODEL STABILITY

Similar to Session 5.2.4, we provide below the HKMA's expectation to model stability for VPBank's reference. VPBank can adjust its own expectation according to its capability, practical conditions, and especially SBV's expectation if it is different to HKMA.

The HKMA expects banks to demonstrate that their internal rating systems exhibit stable discriminatory power. Therefore, in addition to in-sample validation, banks should be able to demonstrate their internal rating systems' discriminatory power on an out-of-sample and out-of-time basis. This is to ensure that the discriminatory power is stable on datasets that are cross-sectionally or temporarily independent of, but structurally similar (i.e. obligors' key characteristics such as industry and company size, in the independent dataset for validation are similar to those in the development dataset) to, the development dataset. If out-of-sample and out-of-time validations cannot be conducted due to data constraints, banks will be expected to employ statistical techniques such as k-fold cross validation or bootstrapping for this purpose. When a bank uses these statistical techniques, it needs to demonstrate the rationale and the appropriateness of the chosen techniques, and understand the limitation, if any, of these techniques.

The HKMA expects banks to establish internal standards for assessing the discriminatory power of their internal rating systems. Breaches of these standards, together with the associated responses, should be fully documented. The HKMA will expect to see a range of responses from increase in validation frequency to redevelopment of the internal rating systems, depending on the results of the assessments.

The HKMA will expect a bank's internal standards for its rating systems' discriminatory power, and its responses to breaches of these standards, to be commensurate with the potential impact on the bank's financial soundness of a failure of its internal rating systems to discriminate adequately between defaulting and non-defaulting borrowers. In setting its standards and determining the response to a breach of those standards, a bank should take into account factors including, but not limited to, the relative sizes of the portfolios to which the internal rating systems are applied, its risk appetite relating to the portfolios, and the inherent risk characteristics of the portfolios.

The main body of the report is already finished, the rest includes the Appendix for the data dictionary for the financial/income statement and CIC information.

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

6. APPENDIX

APPENDIX A. DATA DICTIONARY FOR FINANCIAL/INCOME STATEMENT

Financial Statement

CODE (CT)	FIELD NAME
	Asset
100	**A. Short-term asset**
110	**I. Cash and Cash equivalents**
111	1. Cash
112	2. Cash equivalents
120	**II. Short-term financial investment**
121	1. Short-term investment
129	2. Provision for decrease in Short-term investment
130	**III. Short-term receivables**
131	1. Receivables from customers
132	2. Prepayment to sellers
133	3. Short-term internal receivables
134	4. Receivables on progress of the contract
135	5. Other Receivables
139	6. Provision for receivables
140	**IV. Inventory**
141	1. Inventory
149	2. Provision for decrease in price of inventory
150	**V. Other short-term asset**
151	1. Short term Accruals
152	2. Discounted VAT
154	3. Tax and fees receivables
158	4. Others
200	**B. Long-term asset**
210	**I- Long-term receivables**
211	1. Receivables from customers
212	2. Business capital at subordinates
213	3. Long term internal receivables
218	4. Others
219	5. Provision for long term receivables
220	**II. Fixed assets**
221	1. Tangible fixed assets
222	Original cost
223	Accumulated depreciation
224	2. financial leasing asset
225	Original cost
226	Accumulated depreciation

CODE (CT)	FIELD NAME
	Equity and Liabilities
300	**A. Liabilities**
310	**I. Current Liabilities**
311	1. Short-term borrowings
312	2. Accounts payables
313	3. Prepaid accounts receivables
314	4. Tax payable
315	5. Staff expense payable
316	6. Accrued expenses
317	7. Intercompany payables
318	8. Payables according to the progress of the contract
319	9. Other current liabilities
320	10. Provision for current liabilities
323	**11. Welfare Fund**
330	**II. Non-current Liabilities**
331	1. Long-term accounts payables
332	2. Long-term intercompany payables
333	3. Other long-term payables
334	4. Long-term borrowings
335	5. Deferred tax returns
336	6. Redundancy Allowance
337	7.Provision for long-term liabilities
338	**8. Unearned Revenues**
339	**9. Research & Development Fund**
400	**B. Equity**
410	**I. Equity**
411	1. Invested Capital
412	2. Surplus Equity
413	3. Other reserves
414	4. Treasury shares
415	5. Revaluation reserves
416	6. Foreign Exchange differences
417	7. Investment Development Fund
418	8. Financial Fund reserves
419	9. Other Funds
420	10. Undistributed profits after tax
421	11. Capital Construction Investment

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

227	3. Intangible fixed assets	**422**	12. Business Arrangement Supporting Fund
228	Original cost	**430**	**II. Funds and Other Funds**
229	Accumulated depreciation	**431**	Fund for rewarding and welfare
230	4. WIP basic construction cost	**432**	1. Funds
240	**III. Invested real estates**	**433**	2. Funds invested in Fixed Assets
241	Original cost	**439**	**C. Minority Interest**
242	Accumulated depreciation	**440**	**TOTAL EQUITY**
250	**IV. Long-term financial investment**		
251	1. Investment in subsidiaries		
252	2. Investment in associates and joint ventures		
258	3. Other long-term investment		
259	4. Provision for decrease in long-term financial investment		
260	**V. Other long-term asset**		
261	1. Long-term prepayment expenses		
262	2. Deferred income tax		
268	3. Other long-term asset		
269	**4. Commercial advantages**		
270	**Total Assets**		

Income Statement

CODE (CT)	FIELD NAME
1	**1. Revenues from goods and services**
2	2. Revenue deductible items
10	**3. Net Revenues from goods and services**
11	4. Cost of Sales
20	**5. Gross Profits**
21	**6. Financial Revenues**
22	7. Financial Expenses
23	- Interest Expense
24	8. Selling Expense
25	9. Management and Administrative Expense
30	**11. Net profits from operation**
31	**12. Other Income**
32	13. Other Expenses
40	**14. Other Profits**
50	**16. Profits before Tax**
51	17. Corporate Tax
52	18. Deferred Corporate Tax
60	**19. Profit after Corporate Tax**
70	**20. Basic Earnings per share**

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential

APPENDIX B. DATA DICTIONARY FOR CIC INFORMATION

CODE	DESCRIPTION
CIC1	Months of oldest bank contract
CIC2	Month of youngest bank contract
CIC3	Inquiry times in last 12 months
CIC4	Inquiry times in last 6months
CIC5	Inquiry times in last 3months
CIC6	Q1: When was the borrowing company founded? (day/month/year)
CIC7	Q16: What is the number of lenders outstanding from the CIC form?
CIC8	Q17: How much is the company's total outstanding lending balance from CIC form?
CIC9	Q18: Has the borrowing company had overdue debts in any credit institutions in the last 12 months? (DPD more than 90 days) YES/NO/NA
CIC10	Geography Information
CIC11	Has the customer at observation point has debts in group 2 or not?
CIC12	How many times a customer have debts in group 2 in the past 12 months from observation point?
CIC13	How many months since the latest debts in group 2 the customer have?

Copyright © 2014 EntroFine. All Rights Reserved. Strictly confidential