CHAPTER 11

Summary of Conclusion and Recommendation and Direction for Future Research

11.1 Conclusion and Recommendation

11.2 Direction for Future Research
Chapter 11

Summary of Conclusions & Recommendation and Direction for Future Research

11.1 Conclusions and Recommendation

In this research, effort has been made to identify the weaknesses in credit appraisal as well as credit monitoring and to examine whether the weaknesses has led to NPA. Consequently, the research has tested the effectiveness of the credit appraisal models of ten banks, their monitoring systems, follow-up action and other aspects of credit risk management. Based on the study of 181 cases of NPA and non-NPA status in details within the framework of Basel II requirements, the research has tentatively arrived at the following conclusions. The findings should be considered tentative as the models are in the process of evolution and probably it would be perfected over time, through the application to real-life situation by the respective banks. Further only 181 cases from the North-East has been considered which is not adequate to test the model with so many parameters.

The following conclusions have been drawn:

- Banks have been trying to develop good models considering more and more parameters. The study, however, found that most of the parameters are not effective in the multivariate situation though they are effective in the univariate situation in most cases and few of them are counter-intuitive. In fact, many parameters are highly correlated making the selection of parameters a complicated process. The explanatory power of the models is also not satisfactory and most of the variation in the current status (i.e., default or non-default) is probably better explained by the parameters not included in the model. Hence, there is a need to experiment with new parameters not considered so far by any model.

- The study found non-financial parameters as better default predictors than financial parameters. Though many researchers (Hayden, 2003; Fernandes, 2005; Frerichs & Wahrenburg, 2003) try to suggest internal models based on quantitative financial parameters, we believe that inclusion of qualitative factors into their models will definitely improve the quality of the model. The same
argument has been put forward by Grunert, Norden & Weber (2003). The banks have been trying to develop multiple models to deal with different situations based on nature of activity (e.g., trading or manufacturing), nature of accounts (e.g., existing or new), and type of facility (e.g., working capital or term loan). Most of the models are not effective to the extent it should be. In some cases, parameters are same but weights are different for different situations. In few cases, parameters are different for different situations.

If the performance of a company is more responsive to the industry performance than to the performance of the economy, it is likely that a model accommodating different variables for the different industry sectors will perform better than a model which forces the same variables and parameters to all firms across industries. Fernandes (2003) has also put forward similar views. Hence, it is desirable that banks develop a few industry-specific models where applicable. Few private sector banks are also working in a similar line.

- Thus most of the models have failed to deliver the outcomes as expected. Though they are useful, in most cases, they are either poorly predicting the health of the firm or misclassifying beyond the expected limit. Further models have exhibited weaknesses of various forms, which may lead to adverse selection of proposals. Banks need to refine their models by removing the weaknesses of various forms and also to deal with different situations.

It should be noted that sound model alone cannot serve the purpose as the models give mechanical score whatever may be the level of perfection of the model. Overall human judgement on acceptance of proposals is equally important and the models should leave adequate scope for incorporating human judgement into the score or in the final selection.

- Regarding the results of the analysis of the transition matrix of each bank, none of the banks produce a clearly more attractive transition matrix. If results based on 181 cases are at all indicative of class-wise actual default rates of the entire population of the bank's loan portfolio then estimation of PDs based on historical default rates may not produce the desirable result.

- The number of categories/grades in the risk rating models of the sample banks varied from six to nine in case of the non-default category. Only two banks had defined default classes and have three classes in this category. Other banks are silent on the default categories/grades. The internal rating systems of commercial
banks in the US typically have six to ten rating classes (Krahnen & Weber, 2001). Banks may define two to three default classes to reflect the following new reference definition of default given by Basel Committee on Banking Supervision.

“A default is considered to have occurred with regard to a particular obligor when one or more of the following events have taken place:

➢ It is determined that the obligor is unlikely to pay its debt obligations (principal, interest, or fees) in full;
➢ A credit loss event associated with any obligation of the obligor, such as charge-off, specific provision, or distressed restructuring involving the forgiveness or postponement of principal, interest, or fees;
➢ The obligor is past due more than 90 days on any credit obligation;
➢ The obligor has filed for bankruptcy or similar protection from creditors”.

• Lapses in monitoring is seen in both NPA and non-NPA cases and the lapses visible in the Inspection Reports no way signals the quality of the account and health of the borrowers. This, however, does not mean monitoring is not important. Though the study has not directly explore it has argued that the lapses which are not unearth by the inspectors may have the signaling power and the banks need to look for additional area like financial and strategic standing of the firm through financial and strategic indicators. Simply, stricter monitoring of the parameters indicated in the minimum monitoring requirements of the bank, may not lead to prevention of advances from becoming non-performing.

• Personal visits or Personal Contacts and Compromise have been found to be the most effective recovery measures. Both are interrelated and personal visits lead to compromise, which in turn leads to higher recovery. In view of this finding, banks should intensify the visits with the intention of bringing the borrowers to compromise on the amount of loan to be paid. This strategy is found to be much better than resorting to legal measures.

• NPA recovery has been treated as a criterion for performance appraisal of the branch manager as well as the credit officer but the weightage given is not sufficient enough to motivate the manager to put sincere effort in NPA recovery and is not consistent with the importance NPA has on bank’s profitability,
survivability, fund deployment and interest spread. Hence, the NPA recovery as a criterion of performance should receive more weightage.

- By following the RBI prudential norms, though banks manage the portfolio risk and avoid credit concentration to a particular sector, this may not lead to diversification and reduction of risk of the portfolio unless diversification is across sectors with independent or negatively correlated risk. Regarding risk reduction through diversification, few suggestions have been made in Chapter 7 (Section 7.6), which banks may incorporate in their loan policy.

- An attempt has been made to suggest the required specification of the model rather than designing a specific credit risk model which would serve as a guideline for evolving new models for different situations.

### 11.2 Direction for future research

Some of the areas, which may be good avenues for further research, are highlighted below:

- In this research study the rating models are tested based on 181 observations only. Most of the research in the area have considered large sample of data but taken only the financial part of the evaluation process ignoring the non-financial part. Unlike financial parameters, detailed published data on non-financial parameters are hard to come by. However, one central requirement for banks, to adopt the IRB approach under Basel II, is inclusion of not only traditional quantitative, financial factors but also qualitative, non-financial information. Also since the non-financial parameters have the most discriminating and predictive power, this research has considered the model in totality incorporating both financial and non-financial parameters. Considering the extreme difficulty in creating and collecting data on non-financial parameters (since the model was not in place at the time of initial evaluation of the loan proposal) the size of the sample is low making the results of the study difficult to generalize. The future research work should test the models using a larger samples.

- Because of the restricted number in sample size only nineteen models of the ten banks could be tested. Future research work should also consider other models not taken up in this research study.
The sample covers only one geographical area where default rates are relatively higher. Hence there is a need to test the model considering samples from other geographical areas.

The research has combined both discrete and continuous scales. However, Mählmann (2004) argued that incorporating qualitative factors into a scoring function requires the transformation of nominal or ordinal-scaled data to a metric scale. A standard method of doing this is to use a dummy (0/1) variable for each different category of the qualitative factor, except for a reference category. The dummy takes the value 1 only if the realization of the qualitative factor equals the specified category. Obviously dummy variables are not normally distributed and so they violate the central assumption of classical discriminant analysis. Therefore, Blochowitz & Eigermann (2000) proposed the Fisher-Lancaster scaling procedure for handling qualitative data in discriminant analysis. Though, Mählmann did not find, in his study, any advantage to scaling qualitative factors with the Fisher-Lancaster (FL) method instead of using dummy variables in linear discriminant analysis, considering the need for consistent scaling, future research work should also test whether the scoring model will show better discriminatory power if rescaling of the qualitative factors is done in the line of the FL method.

Kraft, Kroisandt and Muller (2004) has put forth the argument that for evaluating existing and alternative rating systems, we would actually require the full database of all past credit applicants. They have strongly recommended this, as their study found that censoring (considering only accepted credit applicants) leads to biased estimates for any kind of performance measure for rating systems. Like others, this research work has considered only loan proposals, which have been sanctioned, and thereby fails to meet the requirement of the study by Kraft, et al. The model could have been better built had the in-sample database used the loan proposals, which are rejected by the bank. Hence, future research work may look into the model building process considering both accepted and rejected proposals and test whether inclusion of rejected proposals in the in-sample database improves discriminating and predictive power of the designed model. As banks do not maintain a database of the rejected proposals, this type of research will require collaborative efforts between the researcher and the bank on a long-term basis. Although data collection is costly, many banks have recognized its
As discussed in the Chapter 2, there are many techniques for predicting default risk with varying degree of success. None of the techniques can be certified as superior to other techniques in all the situations. However traditionally logistic regression is preferred when outcome is binary. This research also tried to use multiple logistic regression along with multiple regression and discriminant analysis. But since the number of variables considered in the test of the models is high in relation to the number of cases considered, the optimal solution was not found in most of the runs when multivariate logistic regression is used. However, we used univariate logistic regression to test the parameters used in the models where optimal solution was obtained. This study has used MDA and multiple regression for evaluating models. Though the use of multiple logistic regression could have given additional information we do not expect too much from its use as the researchers like Mühlmann (2004) concludes in their study that the difference between logistic regression and linear discriminant analysis is not large enough to favour one technique over the other if the task is classification of firms into only a few groups. However, if a study uses large database, the application of multiple logistic regression and other techniques could be a good avenue for future research.

The current practice of the banks is to convert the score into rating classes for which pre-determined probabilities of default (PD) are available or which maybe calculated from historical data. Future research work should explore whether obtained score may be directly translated into default probabilities using statistical techniques like LR. This is consistent with the approaches that Basel (2001) suggests for PD estimation. Further, it may be tested whether LR default probability estimates provide better accuracy of default estimates than the PDs derived from historical data or external agency data. Frerichs & Wahrenburg (2003) tested the same but results arrived still remain inconclusive.

Hayden (2003) has tested whether the structure and the performance of credit scoring models is sensitive to the default definitions that are used to derive them using only financial information. Future research can test whether scoring models based on other than exclusively accounting information and derived with different methodologies (e.g., neural networks) are also sensitive with respect to the
> Basel (2001) suggests banks to estimate the long-run average PD over an entire economic cycle. In a similar vein, an interesting but lesser probed area for future research could be of testing the sensitivity of models to changing economic conditions i.e., to examine the performance of models across a range of economic environments and cycles. Sobehart, Keenan & Stein (2000) has suggested this for model validation.

> This study tested the effectiveness of a few models of ten public sector banks only. Like the PSBs all the private sector and foreign banks are also developing and using internal rating model to meet the Basel II requirements. The comparison of effectiveness of the internal rating models of public, private sector and foreign banks could be a very interesting research work.

> Last but not least effort to develop neuro-fuzzy credit rating model could be another interesting research area where some initial work has already been initiated in the west. Further, the hybrid of genetic algorithm with neuro-fuzzy could be another area where lot of research work need to be initiated.