

Statistical modeling and analysis of defaulted
and normal states loan's expected loss with
Japanese regional banks' integrated database

TANOUE YUTA

Doctor of Philosophy

Department of Statistical Science
School of Multidisciplinary Sciences
SOKENDAI (The Graduate University for
Advanced Studies)

Statistical modeling and
analysis of defaulted and
normal states loan's
expected loss with Japanese
regional banks' integrated
database

Yuta Tanoue

Department of Statistical Science
School of Multidisciplinary Sciences
SOKENDAI

Abstract

This thesis composed of five Chapters aims contributing to improvements of risk modeling of bank loans with my integrated database.

Chapter 1 gives backgrounds and outline of this thesis. As required by Basel III, banks are allowed to estimate their portfolio's credit risk. To estimate accurate credit risk, detailed analysis of credit risk (Probability of default(PD), Loss given default(LGD), Exposure at default(EAD)) and development of high accurate estimation models are required. However, because of the shortage of publicly accessible database, especially database containing loss information, credit risk has not been much researched. Further, although there are studies using not publicly accessible database, most of these database contains only limited information and composed of only single bank. Further, there are no publicly accessible bank's loan database that can be used for credit risk research in Japan. To acquire more general results and knowledge, I developed an integrated database composed of five Japanese regional banks' loan data, which contains a great deal of information about factors that might influence the credit risk. With my integrated credit risk database, I tackled three issues in Chapter 2-4.

Chapter 2 developed a model that estimates the probability of return-to-normal-state from default loan to performing loan with logistic regression model to identify the significant factors on return-to-normal-state. There are many researches on credit risk of the non-default state loans. On the other hand, there are only a few researches which studies on the credit risk of the default state loans. Studying on credit risk of default state loans is important because default state loans also can cause credit risk. I used corporate loan dataset, which contains information on borrower characteristics, financial characteristics, and time period categorical variables as explanatory variables of the proposed model, from a regional bank recorded every half year from September 2007 to March 2012. I propose a YJ transformation for the explanatory variables, which is an expansion of Box-Cox transformation to negative values. As a result, our result confirmed that although YJ-transformation do not statistically significantly improve model predictive performance, this transformation can improve model predictive performance. Furthermore, I found that debtors with high neglog-transformed number of officers and employees at end of term, model score, YJ-transformed Deposit collateral coverage/EAD, YJ-transformed Land/Total asset and YJ-transformed Notes payable/Total asset are likely to return-to-normal-state. On the other hand, debtors with high YJ-transformed Long-term working capital/EAD are unlikely to return-to-normal-state.

In Chapter 3, we analyze the effect lending area on regional banks portfolio's credit risk. To develop accurate bank loan's credit risk estimation model, we have to analyze the influencing factors. Existing studies consider the effect of macro economic, global factors, country factors, industry factors and so on. The depopulation in provincial areas causes the decline of these area's economies in recent years in Japan. Accordingly, Japanese regional banks have been advancing into areas outside their original operational bases, especially into major cities and urban regions. Since amount of lending to areas outside the original operational base is increasing and such lendings's uncertainty is being high, the lendings are expected to significantly affect the credit risk of the regional banks' lending portfolios. Therefore, we analyze the effects of the

lending areas on the credit risk of the regional banks' lending portfolios. Using five regional bank's credit data, we describe the fundamental statistics of each area, we develop credit risk estimation models to analyze the effect of lending area variables. These analysis indicate that lending area variables affect the credit risk, default is less likely to occur outside the local region and in Tokyo, and LGD is small in the local region as compared as outside the local region and in Tokyo.

In Chapter 4, using a dataset of five Japanese regional banks, we propose LGD estimation models that combine a two-stage model with classification and regression tree (CART)-based boosting and SVR with a logistic transformation. We developed (I) Loss occurrence model of two-stage model with Platt scaling model in which we use the output of CART-based Boosting as the predictor and the second component is developed with SVR with a logistic transformation. As a result, we confirm that our model is superior to other models, indicating that a combination of the two-stage model with CART-based boosting and SVR with a logistic transformation works well for LGD estimation.

Chapter 5 describes the conclusion.