Application of Discriminant Analysis of Home Mortgage Loan Prepayment Risk

Cheng-Chung Wu¹, Wang-Te Wu², Shu-Man Tsai³, Min-Chih Lu⁴

¹Assistant Professor, Department of Administration, Cheng Shiu University
²Lecturer, Department of Administration, Cheng Shiu University
³Master Student, Department of Public Finance and Taxation, National Kaohsiung University of Applied Sciences
⁴Master Student, Department of Industrial Engineering Management, National Kaohsiung University of Applied Sciences

Abstract

Whether the borrower is engaging in default risk behavior is bound to be driven by some incentives as judged by internal and external observable variables; or it can be observed from some of the characteristics of the borrower. Which type of borrowers is less reliable in credit and may prepayment is an issue of great concern to asset risk management of financial institutions holding the housing mortgage loans. Therefore, the study of the tracking and forecasting of changes in relevant factors during the loan period of each borrower is very valuable. For example, the prediction and evaluation of financial institution prepayment risk probability by changes in borrowers’ income can help prevent the borrowers’ prepayment risk behavior in advance.

Keyword
Discriminant Analysis; Home Mortgage Loan; Prepayment Risk

Introduction

Prepayment probability is a risk to real estate mortgages as they both affect the value of creditors’ rights. There must be incentives that drive borrowers to be engaged in these two risky behaviours. The profiling of certain borrowers can shed light to which type of borrowers is more able to make prepayment or simply less reliable (hence likely to default). These are all the issues that financial institutions should take into careful consideration in risk management of mortgage businesses. Meanwhile, in order to establish an understanding of the characteristics of prepayment and overdue repayments across borrowers, the analysis of the historical data of mortgages offered by the banking industry in Taiwan is the first priority and essential.

Literature Review

Steenackers and Goovaerts (1989) used stepwise logistic regressions to establish Credit Scoring Model of personal loans. Their research found that the ages, ownership of phones, duration of staying in current addresses, locations, occupations, public servants or not, monthly salaries, ownership of the residence, the number of previous mortgage undertaken and years of borrowing all have significant influences on the normal repayments going forward.

In fact, the decisions of borrowers should be based on the timing of the decision making. They make the decision to default or make prepayments based on their assessment of their own conditions and economic situations. Basically, some factors that influence the risk behaviours of the borrowers do not change over time, such as sex, purchase prices and repayment methods. However, many factors do change over time, such as the contractual rates of FRMs, present values and ages of the real estates, or even trigger events such as unemployment rates.

Therefore, in the consideration for prepayments, borrowers will gauge whether the contractual rates are high or low and whether prepayments are appropriate (Chiang, 2007). In the consideration for defaults, they will examine whether the present value of the real estate is equivalent with the balance of the remainder of the mortgage. If the housing prices have been plummeting to a very low level, borrowers will consider whether it is still necessary to keep repayments. As to the purchase prices and the original low contractual rates may be things of the past and
falling outside the scope of consideration.

**Research Method**

The main purpose of this study is to compare different prediction models for defaults and prepayments of mortgage loans. The key is to identify a better mix of predictive variables. Therefore, this paper analyzes the mortgage data of a Taiwanese bank over the past 6 years. There are a total of 16,213 entries of sample data. The sample is highly representative.

This paper compares different prediction models for defaults and prepayments of mortgage loans so as to identify a better mix of predictive variables (rather than a better analysis method). Therefore, this paper only uses the traditional Discriminant Analysis to analyze the mix of predictive variables.

Discriminant Analysis aims to identify a linear combination of independent variables and discriminate them into groups based on the linear discriminant functions, so as to maximize the inter-group variances in relation to within-group variances and ensure the best discriminant effects. As many other multivariate statistic methods, discriminant analysis is established on the assumption of linear statistical models and observation values. Therefore, the coefficient values of the discriminant functions can explain the results of discriminant analysis. There are two methods of discriminant analysis, i.e. parameter method and non-parameter method. This paper uses the parameter method for group discriminations. Like multiple regressions, discriminant analysis is the most frequently used classification method. The key difference between these two methods is the attributes of independent variables. The independent variables of multiple regressions are mostly continuous; whereas the independent variables in discriminant analysis are ordinal variables or class variables. Discriminant analysis groups data in according with data characteristics.

**Discriminant Analysis of Prepayment Behavior Model**

The early-repayment behavior model also separates males and females into two groups. A separately constructed model is used to perform the discriminant analysis of the early-repayment behaviors.

An overall discriminant analysis on the 16,213 entries of data is performed. Approximately 61.8% of the borrowers who made prepayment can be correctly classified by the discriminant functions. The result is rather significant. In the cross reliability analysis, 61.7% of the borrowers who made prepayment and 51.3% of the normal ones can be correctly classified by the discriminant functions. As far as the cross reliability analysis is concerned, 51.2% of the normal borrowers can be correctly classified by the discriminant functions. Therefore, the overall accuracy of the discriminant analysis is 57.9%, and that of the cross reliability analysis is 57.8%. This indicates that as far as the overall data of borrowers is concerned, the discriminant analysis will exhibit significant judgment ability over the borrowers who make prepayment. A comparison of the eigenvalue of 1 for the discriminant function and the value of 0 for the discriminant function shows that it has a better discerning ability. Therefore, it can be used to perform discriminant analysis of prepayment based on the variable characteristics. According to the overall discriminant analysis of the prepayment, Wilks’ Lambda value of 0.972 shows that the significance of the function 1 reaches p-value<0.000 and carries very good explanatory power. Table 6 shows that explanatory power coefficients of discriminant functions for prepayment.

This paper establishes the discriminant functions on prepayment on the basis of the characteristics of overall variables. The discriminant functions of 1 are as follows:

\[
Z_{\text{prepayment}} = 0.574 \text{ loan at the beginning} + 0.311 \text{ outstanding balance} + 0.322 \text{ financial situation} + 0.372 \text{ grace period for payment} + 0.348 \text{ percentage of leverage} - 0.751 \text{ loan offered} + 0.132 \text{ appraisal value} + 0.578 \text{ No. of defaults} - 0.311 \text{ appraisal value}
\]

According to the discriminant functions, the variables that have positive influence on prepayment are the number of defaults, loan at the beginning, mortgage duration, the percentage of leverage, financial situation and appraisal values. In fact, the higher the number of defaults, the more influence they will have on prepayment.

Of course, the financial situation, loan offered and mortgage duration are also in the equation. The higher the loan
offered at the beginning and the longer the mortgage duration, the more likely prepayment will occur.

<table>
<thead>
<tr>
<th>Function</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan at the beginning</td>
<td>+</td>
</tr>
<tr>
<td>Financial situation</td>
<td>+</td>
</tr>
<tr>
<td>Grace period</td>
<td>-</td>
</tr>
<tr>
<td>Mortgage duration</td>
<td>+</td>
</tr>
<tr>
<td>Percentage of leverage</td>
<td>+</td>
</tr>
<tr>
<td>Loan offered</td>
<td>+</td>
</tr>
<tr>
<td>Appraisal value</td>
<td>+</td>
</tr>
<tr>
<td>No. of defaults</td>
<td>+</td>
</tr>
<tr>
<td>Age of the borrower</td>
<td>-</td>
</tr>
</tbody>
</table>

The variables that have negative influence on prepayment are loans offered, grace period and age of the borrowers. This paper also finds that the lower the actual amounts borrowed (loans offered), grace periods and age of the borrower, the more likely the prepayment will occur. This also explains that the higher the borrowed amounts and number of defaults, the more likely the borrowers will switch their mortgages. In case of defaults, the creditors (banks) will also collect the overdue payments from the default borrowers. Under this situation, the borrowers with better financial situations are more likely to make prepayment.

Conclusions

This paper has examined past studies on prepayment of mortgage loans by analyzing the introduction of variables and definitions. It compares the prediction models of prepayment of mortgage borrowers in order to identify a better mix of predictive variables (rather than a better analysis method). Therefore, this paper only uses the traditional Discriminant Analysis to analyze the mix of predictive variables.

This is consistent with the analysis of the prepayment by male mortgage borrowers. According to the discriminant function in the constructed discriminant analysis model, this paper finds that as far as the significance of the individual variables are concerned, the discriminant functions will exhibit significant effects in the analysis of prepayment. The variables that influence the level of mortgage sizes are also rather significant. In the model where males and females are separated in the data, it is inferred that the changes in the appraisal value is the main contributor to the variance of these two functions.

REFERENCES