



Applying Fisher Discrimination Approach to Assessing Customers' Risk in Bank Card Business

Yongju Zhang & Erhong Hou

Business Department

Qingdao College Ocean University of China

Qingdao 266300, China

Tel: 86-532-8999-3151 E-mail: zhangyongju@yahoo.com.cn

Abstract

The development of bank card business has severely challenged the improvement of risk-controlling technology. While assessing risk is the foundation of risk-controlling in bank card business. Presently, domestic banks have set up their customer data base. By analyzing the status of risk assessing in risk-controlling, this article tries to applying Fisher Discrimination Approach to assessing customers' risk and overcomes the difficulty of determining the influential factors' share in traditional bank assessing approach.

Keywords: Fisher Discrimination Approach, Bank card, Risk assessing

As a high-profit product, bank card has become the very important income resource of service business in foreign banks. In recent years, the income contribution rate of service business in large international banks has reached 30% -70%. Among them, Citibank has reached 70%. Britain's Barclays Bank is about 73%. Till to 2007, Citibank reaches 80%. However, the service business of commercial banks in China accounts for a proportion of average net income only 14%. Just due to the high profitability and strong expansion space, combined with the recent liberal policies on individual consumption credit, China's commercial banks begin to exploit bank card business. However, a variety of reasons cause it' s development to be hampered, such as many marketing methods without norms, the lack of a sound domestic personal credit system, as well as the lack of experience and matched risk-control techniques. Chinese commercial banks' bank card business annually suffers from important loss. Analyzing the causes of non-performing assets' forming, this loss has greatly challenged the improvement of bank card business' risk prevention and controlling technology from different perspectives although their risk exposures are very different from each other.

1. The significant of assessing customers' credit risk for controlling bank card business' risk

From the point of view of products features, in addition to the general features of bank products' risk, bank card's risk also has some different characteristics, such as the plastic carrier, the digitized form of transaction, business quickly accepted, unsecured loan, quota credit and so on. Therefore, assessing customers' credit risk will accompany the whole process of using bank card.

Due to economic rapid development, bank card holder's social composition also changes all the time. Changes of income result in changes of paying ability. Someone overdrafts maliciously and crimes by cheating. Many kinds of objective situation and subjective moral factors make card holder reject paying. This has highlighted assessing customers' credit risk and monitoring in a timely manner.

In term of banks' running, as an enterprise, banks take profit maximization as their ultimate goal. According to the positive correlation of risk and revenue income, to deal with bank card's risk should follow the principle of maximizing revenue income with an acceptable level of risk, rather than simply seeking to minimize the loss or zero risk. This requires that banks should find the best balance or combination way among risks, revenue and efficiency, so that they can achieve a balance between bank card's security and convenience which is the very important premise for products' vitality and sustainable development. At the appropriate risk, the capacity and the technology of controlling risks become the key to banks' profitability.

Because of the fierce market competition, the common problem currently faced by card issuers is how to realize the double win between the issuer amount and risk control. Assessing customers' risk is the foundation during the whole process of risk control including risk identification, risk estimation and risk assessment. Only by assessing customers' risk and issuing cards to the applicants, bank card business can begin to run. Accurately assessing customers' risk is an important basis for quota credit. It can play a role of early warning. Based on it, bank can develop appropriate risk

control policies so as to improve its statement and revenue level. As a result, assessing customers' risk plays a very important role throughout controlling bank card business' risk.

2. The introduction and the basic principles of Fisher Discrimination Approach

The traditional approach of assessing customers' risk in Chinese commercial banks is credit score. By comparing the applicant or the borrower's credit history with all borrowers' information in the database, banks check whether the applicant's development trend of statement are the same with those who regularly breach of contract and even bankrupt so as to fall into financial difficult situation. Mainly relying on customers' past credit information which include advantages and disadvantages, this approach gives different index different values and weights them. There are many kinds of calculation methods on credit score. The best-known in the world is FICO credit scoring model which is invented by U.S. engineer Bill Fair and mathematician Earl Isaac in the 1950s. It is still used by three major credit agencies of the United States. But the traditional method has defects on quantifying risk factors which impact bank card and on determining the weight of factors. While FICO credit scoring model is based on that the United States has a sound individual credit system which has not yet owned by China's commercial banks. If it is blind introduced, it can not effectively control bank card risks.

In the 1930s, Fisher presents an approach on linear classification. According to the different types of targets, it adopts appropriate discrimination factors and establishes the relationship formula on factors and targets of discrimination. By selecting the appropriate rules of discrimination and judging a number of factors, it classifies the observed samples. This approach has been widely applied to all fields of medical identification, chemical analysis, rating natural disasters, image recognition, rating credit risk of the stock market listed companies, and so on.

The basic idea of Fisher Discrimination Approach is that: project the points in N-dimensional space to one direction. After transferring data, put the same category points together as far as possible. Meanwhile separate the different points as far as possible. So one can only analyze data in one-dimensional space and classify. This approach does not need the probability distribution of samples. It also does not restrict the independence of variables. However, the probability distribution of samples and the independence of variables are difficult to be achieved during assessing customers' credit risk. Therefore, it is appropriate for this paper.

3. The calculating process of Fisher Discrimination Approach

After selecting one direction, Fisher Discrimination Approach projects the points in N-dimensional space to one-dimensional space. It makes the variance of the same category points as small as possible but the different points' as large as possible. Thus, the values of different samples' risk assessment will have a unified standard.

Assuming that the risk level of all history samples is divided into m, $R^i (i=1,2,\dots,m)$ stands for each level of risk. Symbol “ \succ ” said high or low of different risk levels. For example, “ $C^i \succ C^j$ ” said that the level of i-type risk is higher than that of j-type risk. Higher risk level accompanies lower credit rating. We can get: $C^1 \succ C^2 \succ C^3 \dots C^{m-1} \succ C^m$

If the number of samples belonging to $C^i (i=1,2,\dots,m)$ is $n_i (i=1,2,\dots,m)$, each sample is explained by k index which have the same attributes. Thus, one sample which belongs to $C^i (i=1,2,\dots,m)$ reads

$$C_j^i = [C_{1j}^i \quad C_{2j}^i \quad \dots \quad C_{kj}^i]^T \quad (i=1,2,\dots,m; j=1,2,\dots,n_i)$$

Fisher Discrimination Approach aims at seeking a k-dimensional vector $y = [y_1 \quad y_2 \quad \dots \quad y_k]^T$ which projects every sample to one-dimensional space. Z_j^i is the inner product in one-dimensional space, the formula for that:

$$Z_j^i = y^T C_j^i \dots\dots\dots (1)$$

Let $\bar{Z}^i (i=1,2,\dots,m)$ stand for the projection variance of $C^i (i=1,2,\dots,m)$. Write down:

$$\bar{Z}^i = \frac{1}{n} \sum_{j=1}^{n_i} y^T C_j^i = y^T \frac{1}{n} \sum_{j=1}^{n_i} C_j^i$$

Condition 1: if the same category samples' projection variance is as small as possible, that is to meet:

$$\min \sum_{j=1}^{n_i} (Z_j^i - \bar{Z}^i)^2 \quad (i=1,2,\dots,m)$$

Let \bar{Z} be all category samples' projection variance:

$$\bar{Z} = \frac{1}{\sum_{i=1}^m n_i} \sum_{i=1}^m \sum_{j=1}^{n_i} y^T C_j^i = y^T \frac{1}{\sum_{i=1}^m n_i} \sum_{i=1}^m \sum_{j=1}^{n_i} C_j^i$$

Condition 2: while (1) should also meet that the variances of different credit category samples are as large as possible, namely:

$$\max \sum_{i=1}^m \left(Z^i - \bar{Z} \right)^2$$

If formula (1) and the above two conditions are met, one can get the value of y. according to formula (1), a new sample $x = [x_1 \ x_2 \ \dots \ x_k]^T$ multiplies y. get the projection of x in one-dimensional space. Write down: $f(x) = y^T x$. Let $x^* = f(x)$, then classify x according to the distance between x^* and $C^i (i=1,2,\dots,m)$.

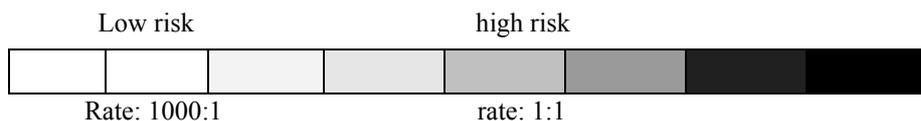
$$\text{If } \left| x^* - \bar{Z}^i \right| = \min \left\{ \left| x^* - \bar{Z}^j \right| \right\} (j=1,2,\dots,m), \text{ x belongs to } C^i.$$

4. the application of Fisher Discrimination Approach to assessing customers' risk

According that the data of applicants or cardholders are not related to each other and the characteristics of individual data are diverse, banks can set up database of assessing customers' risk from the large data warehouse, which is called the target data. By analyzing the characteristics of applicants or cardholders, form a Boolean multi-dimensional customer data resource. Then deal with extraneous noise of these data. The shifted data are the target data that Fisher Discrimination Approach need.

4.1 ascertaining risk levels

The individual sample's credit risk value is ascertaining figure in history data. But the values in the whole database are continuous from low risk to high risk. The customer group can compose a continuous credit value chain. Such as:



According to customers' history notes, credit habit, development trend of statement, the law of breach of contract and financial difficult situation due to bankrupting, ascertain the amount and the critical value of different risks. Namely make sure of the value of m and $C^i (i=1,2,\dots,m)$ in Fisher Discrimination Approach.

4.2 shifting data and indexes

Not all indexes gotten from customers' initial data can be used in measuring risk level. First of all, analyze the correlativity of the data in statistics method so as to dislodge strong correlative data. Meanwhile, deal with extraneous noise of these data to improve the weight of the single index. After shifting data and indexes, get the target data and indexes. They still cannot be used and need shifting. For example, one cannot directly use personal month income, personal disposable income, month expenses, and so on. But should shift them into more integrated index to measure customers' payment ability, such as by $R = (m+p+b+s)/I$. Among them, R is the rate of debt and income; m is mortgage payment; p is installment payment; b is credit card payment; s is credit circulation payment of retailing; I is monthly income.

By shifting them, get the prepared data and indexes. Namely, ascertain $C_j^i = [C_{1j}^i \ C_{2j}^i \ \dots \ C_{kj}^i]^T (i=1,2,\dots,m; j=1,2,\dots,n_i)$. The projection variance in the Discrimination formula can be calculated.

All the above are gotten from history data. According to them, one can make out y in the Discrimination formula.

4.3 testing samples

For any sample $x = [x_1 \ x_2 \ \dots \ x_k]^T$, one can get its $Z_j^i = y^T x$ which meets the two conditions and formula (1). After projecting the sample from N-dimensional space to one-dimensional space, scmpare it with $\left| x^* - \bar{Z}^i \right| = \min \left\{ \left| x^* - \bar{Z}^j \right| \right\} (j=1,2,\dots,m)$.

According to $C^i (i=1,2,\dots,m)$, ascertain x's risk level.

4.4 verifying Fisher Discrimination Approach

The validity of the formula coefficient y achieved from the above process need verifying. However, due to Chinese

commercial banks' exploiting their credit card business only from 2003, together with the incomplete social credit system, the existing database is not large enough to verify the formula by a part of data which are put away. Based on the actual situation in Chinese Commercial Banks, the current database applies to cross-validation approach to verify the validity of the formula coefficient. Banks can put away a part of historical data as verification samples and ensure that they are random. Calculate the validity of the formula coefficient y in the other part of data. After many occasions' repeating this process, achieve a comprehensive amendment. Take it as the final value.

After achieving the result, banks should continue monitoring its validity even if it is effective enough at the beginning. Change the detection method, if necessary. Therefore, a successful model calls for dynamically developing.

With the development of policy and market, the value of indexes and the weight of impact factors applied to assessing customers' credit risk are changing all the time. Banks should often verify the formula and timely update the value of y so as to complete the index system and maintain discrimination effective.

5. The advantages of applying Fisher Discrimination Approach to assessing customers' credit risk

This method combines Fisher Discrimination Approach to assessing customers' credit risk. Based on customers' credit record, relying on a computer powerful computing capability, it provides an objective technology for bank card business to control its risks. And it overcomes the difficulties of traditional credit scoring methods and FICO credit scoring model. It has the following relative advantages:

5.1 The loose conditions of Fisher Discrimination Approach

Whether Logistic regression method, the neural network, classification tree method, or any other linear classification method, scoring models based on them strictly regulate the samples' state and attribute. While Fisher Discrimination Approach requires less conditions. This approach does not need the probability distribution of samples. It also does not restrict the independence of variables. However, the probability distribution of samples and the independence of variables are difficult to be achieved during assessing customers' credit risk.

5.2 The objective and comprehensive index

This credit risk assessing model is based on the principle of large numbers and statistics technology. The indexes in the model are digital, which do not need artificial quantization process. So it avoids the disadvantages of traditional scoring models and reveals the universal law of bank card customers' credit risks. However, the indexes in traditional models are achieved from artificial quantization process.

While preparing for the model, quantities of influential deriving variables are assessed, compared with each other and selected in statistics technology. So the final equation is usually verified by the multi-dimension representative information. It can comprehensively assess customers' credit risk.

5.3 The accurate conclusion and the unified classification standards

No matter how many times repeat the assessment process, only if use this approach, any businessman can draw the same conclusion at any time, in any branch. So the classification standards are unified. Multiplying the coefficient matrix achieved during the first process by the test samples, get the result of scoring. In addition, the model dynamically develops. Banks often verify the formula and timely update the value of y so as to maintain discrimination effective. So it can accurately ascertain customers' credit risk level.

5.4 the high efficiency

This credit risk assessment model can be put into computers. So long as putting into the relevance information, the computer implements automatically and make out the results just within several seconds. So it applies to assessing customers' comprehensive information, especially when customers' data are complicated. However, if only artificially assessing millions of applications' statement and credit risk, the labor costs a lot and accepting speed is too low. Therefore, it has high efficiency.

References

- Chen,Jian. (2005). "Survey on Credit Scoring models". *Chinese Credit Card*.
- Chen,Lin & Zhou,Zongfang. (2004). "Credit Risks Grades Differentiating Based on Fisher Discrimination". *Academic Journal of Applying Basis and Project Section*.
- Chinese Economy Times. (September 14, 2005).
- J.B. Kaoait, E.I. Aiteman & P. Nalayanan. (2001). *the evolution of managing credit risk*. Mechanical industry press.
- Lyn Thomas. (2001). a Survey of Credit and Behavioral Scoring. *Forecasting Financial Risk of Lending to Consumer*. University of Edinburgh Workpaper.
- Zhou,Hongliang & Mu,Wenquan. (2002). *Management of Credit Card Risk*. Chinese Finance Press.