

Research on the Influencing Factors of the Lender's Lending Behavior in China's P2P Network Lending

Zhang Weihong¹, Wang Kecheng², Wang Ximeng³

¹Xidian University, School of Economics and Management, 266, Xi Feng South Road, Xian 710126, China
Whzhang[at]mail.xidian.edu.cn

²Xidian University, School of Economics and Management, 266, Xi Feng South Road, Xian 710126, China
ppboy_rss[at]163.com

³Xidian University, School of Economics and Management, 266, Xi Feng South Road, Xian 710126, China
710757277[at]qq.com

Abstract: *This paper takes the lenders' lending behavior in P2P network lending as the research object, innovatively uses the theory of perceived value to establish the theoretical model of the influence of the related factors on the loan application, uses logistic regression model and two-step cluster method as the analysis tools, to analyze the effect and mechanism of related factors on the success of loan application, then puts forward relevant suggestions on P2P network lending transactions. The study found that: 1. Loan success rate in P2P is low, only 29.65%. 2. Borrowing rates, special type of bid, posting photo, the success and failing times of the loan all have important impacts on the success of loan; the amount of loan has little impact on the success of loan. 3. The factors such as lending credit, posting photo, the success and failing times of the loan, etc. are the most important factors for the clustering results.*

Keyword: P2P network lending, perceived value theory, lending behavior, logistic regression analysis.

JEL: G23

1. Introduction

P2P network lending is a kind of Internet finance services with both the characteristics of small private loan and network lending. Small private loan doesn't need bank and other financial intermediary, will bring together a small amount of personal funds, after the review, and directly lend these funds to borrowers. Professor Mohamed Yunus has made outstanding contributions to the poverty alleviation work in the form of small private loan in Bangladesh and wins the 2006 Nobel Peace Prize, so that this financing model has attracted the attention of the world. The P2P network lending adds wings of network technology to small private loan, which makes it switch from offline to online, so it possesses the characteristics such as convenient, transparent, without geographical restrictions and so on, allowing more people to enjoy microfinance services. In the following years after the 2008 financial crisis, the traditional financial institutions have been hovering in the doldrums, while the P2P network lending use its comparative advantage to come into a rapid development period.

2013 is known as the first year of China's Internet Finance, with the trend of domestic consumption upgrading and a large number of small loan demand bringing by the "popular entrepreneurship and innovation" policy, China's P2P network lending platform has entered a rapid growth period. According to the data of Industrial Bank of China, as of January 1, 2016, China's P2P network lending platforms are more than 2000, including about 350 active trading platforms. As of October 31, 2016, China's P2P network lending industry's historical cumulative turnover reaches 2965.033 billion Yuan, three trillion is close at hand.

In the single month of October, the turnover reached 188.561 billion Yuan, rises by 57.60% year-on-year; China net loan boom index is 127.89 in this month, essentially equals that of the previous few months, and is far higher than 100, suggesting a better development prospects of China's P2P network lending industry.

The P2P network lending is a rapidly developing emerging industry in recent years, but more and more problems are exposed in recent years. From a macro perspective, there are problems of lack of supervision, laws and regulations; from the perspective of P2P network lending industry, there are problems such as lack of canonical credit investigation system, lack of restraint mechanism for borrowers, platform self-financing and other issues. The problems are summed up in two points: first, asymmetric information; second, high risk. Information asymmetry leads to low efficiency of market transaction and low loan success rate; the high risk needs the high risk compensation, the higher the risk of lending, the higher the interest rate demanded by lenders. These two problems seriously affect the transaction enthusiasm of both sides on the P2P network lending platform, thereby hindering the healthy and rapid development of the industry. Studying the mechanism and influencing factors of the lender's lending behavior (whether to bid) is the basis for the government to formulate relevant policies and regulations, but also the basis for the platform to design a reasonable and efficient operation mode. In recent years, foreign scholars mostly use the data of Prosper, Zopa, Kiva these representative platform to analyze, many research results have been published (Iyer, Khwaja, Luttmer, Shue, 2009; Pope, Sydnor, 2008; Ravina, 2007). Scholars explore the relationship between the borrower and the lender in P2P

network lending from the perspectives of economic, information technology, social science and other aspects. The main contents of these studies are focused on the borrower's behavior analysis, the factors affecting the success of the loan, the operation model of the platform, the protection of information and anti-fraud. China's researches of P2P network lending are mainly about the qualitative analysis such as the concept, characteristics and advantages of P2P network lending, and suggestions about regulatory measures and risk control method, but the quantitative analysis on the influencing factors of the lender's lending behavior in china's P2P network lending, is still relatively scarce.

2. Literature Review

2.1 Foreign research state of P2P network lending

The foreign scholars have made broad research about the factors influencing the lending behavior of the P2P network lending, and the following scholars have discussed this issue from different aspects. Spence (1973), Riley (1975), Rothschild and Stiglitz (1976) believe that in an imperfect market environment, the lender can evaluate the credit risk through the analysis of the borrower's financial factors, demographic factors, social capital and other factors, and then make trading decisions. (1) Financial factors. Iyer et al. (2009) study whether they can use other factors to evaluate the credit of the borrower besides the credit rating, the results find that the 28% of credit change from AA level to HR level can be explained by other features, when lenders evaluate borrowers' credit, it is mainly based on standard-banking indices such as debt to income ratio, the amount of arrears of borrowers, the maximum borrowing rate proposed by borrowers^[1]. (2) Demographic factors. Pope (2008) analyzes the relationship between the borrower's age and the success of the loan. The loan success rate of people under the age of 35 outperforms the 35-60 age group by 40-90 basis points. The loan success rate of the 35-60 age group outperforms those over the age of 60 by 1.1%-2.3%^[2]. Laura & Yuliya (2014) find that the borrower's gender, age and other factors such as posting photos will affect the lender's lending behavior; the middle-aged, male borrowers who have posted photos are more likely to borrow money successfully^[3]. (3) Social capital. M. E. Greiner & Wang (2009) believes that social capital is positively correlated with accomplishing the full scale of the loan amount (hereinafter referred to as the Full Scale), can reduce the borrower's loan interest rate, but also can improve the low credit rating borrower's influence, also found that when a borrower borrows money from someone inside his community, he will have a slightly higher loan amount and a lower default rate comparing to someone outside his community^[4]. Collier & Hampshire (2010) find that the bigger group borrower belong to, the lower loan interest rate is, one cause may be in the larger group, each loan application will be faced with more group members' review, so the credit risk is low^[5]. (4) Other factors. Sonenshein & D holakia (2011) find that if a borrower can make a reasonable explanation for their borrowing purposes and make a detailed description of his own situation, it will improve his loan success rate, even if this borrower's credit

rating is low. Klafft (2008) believes that in an anonymous online lending environment, if the lender lacks the experience of loans, will increase the risk of their loans^[6].

2.2 Domestic research state of P2P network lending

Domestic scholars on the research of P2P network lending mainly focus on the research of its model, the introduction of its function and characteristics, the analysis and prevention of loan risk, and the construction of system and regulations.

Lu Xin, Li Huimin (2015) believes that in view of the risk problems in China's P2P network lending, we should clear regulation subject, strengthen the system construction, and build a unified canonical national credit investigation system and other measures to prevent these risk problems^[7]. Liu Hui, Shen Qingjie (2015) through the analysis of five categories of risk in China's P2P network lending, put forward the proposal from the credit investigation system, the disclosure of information, credit rating and industry standards these four aspects^[8]. Wang Changjiang, Yang Jinye (2015) through the analysis of the development mode and trend of China's P2P network lending, believe that we should adopt the mode of "industry self-regulation + government regulation" to effectively monitor the P2P network lending platform in China^[9]. Li Guangming, Zhou Huan (2011) through empirical analysis, extract the key characteristics of defaulting borrowers in P2P network lending, help lenders and P2P network lending platform to control the risk^[10]. Xin Xian (2009) summarizes the operation mode and characteristics of foreign P2P network lending platform, believes that foreign P2P platform can be divided into the following three categories: public welfare type, intermediary type and composite intermediary type which plays the role of intermediary, guarantor, debt collector and interest rate setter^[11].

In summary, we found that financial factors, demographic factors, social capital and other factors are the four core factors that affect the lending behavior of P2P network lending. In addition, the foreign researches on the P2P network lending are relatively mature from both the qualitative and quantitative analyses, but the domestic researches in this field are relatively scarce, especially the empirical researches from the quantitative aspects are more scarce, lack of a scientific and accurate model to quantify the effects of various factors on the lending behavior in China's P2P network lending, which affects the rigor and scientific nature of China's P2P network lending policies and regulations, and is not conducive to judging the transaction risk scientifically and accurately among dealers, thus affecting the long-term healthy development of the industry. We hope that through the empirical analysis of the influencing factors of lenders' lending behavior in P2P network lending, find out the otherness of the different influencing factors, and hope to provide some suggestions to improve the loan success rate for borrowers, and provide some guidance to measure the credit risk for the lender at the same time.

3. Theoretical model and research hypothesis

3.1 Theoretical model

In this paper, the basic assumption is that the lender is rational in the choice of whether to give borrower a loan, as well as the choice of the interest rate. Before the lending behavior, the most important questions is what is the decision-making mechanism of the lender's lending behavior, and what factors will affect the lender's lending behavior decision. On the P2P network lending platform, the borrower first submits the loan request, that is, the loan list, then lender determines whether to bid for borrower's loan request based on the information of loan list. This process is similar to the process for consumers to buy goods, lender is the consumers, loan request is goods, loan list information is goods description, and the lender completes the purchase by bidding activities. In the purchase process, the consumer decides the purchase decision according to the perceived value of the goods. Peter & Tarpey (1975) believes that when consumer is in the purchase process, there are two kinds of perception: consumer perceived reward (desired characteristics of goods) and consumer perceived risk (bad characteristics of goods), perceived reward minus perceived risk is the net perceived reward^[12]. Monroe & Krishnan (1985) propose perceived value theory, which holds that consumers' purchase decisions are affected by the difference of consumer perceived benefits and consumer perceived costs of a product, and this difference is perceived value^[13]. Perceived reward, perceived risk and net perceived reward in Peter & Tarpey's research are perceived benefits, perceived costs and perceived value in Monroe & Krishnan's research. In this paper, we use the concept of perceived benefits, perceived costs and perceived value.

By analyzing the resources that must be paid for the purchase of the goods, the importance, scarcity of goods, and the transaction risk, etc., the consumers form their perceived benefits and perceived costs of the product. Therefore, the acquisition of relevant goods information is the key to make a purchase decision^[14]. In the P2P network lending market, lenders (consumers) through the analysis of the loan list information, form the perceived benefits and perceived costs of loan request (goods), then compare the difference of them and determine their lending behavior (whether to buy).

In this paper, based on the analysis of the foreign and domestic research of the influencing factors of lending behavior, as well as the loan list information in PPDAl, then a theoretical model is established in the light of perceived value theory, as shown in figure 1:

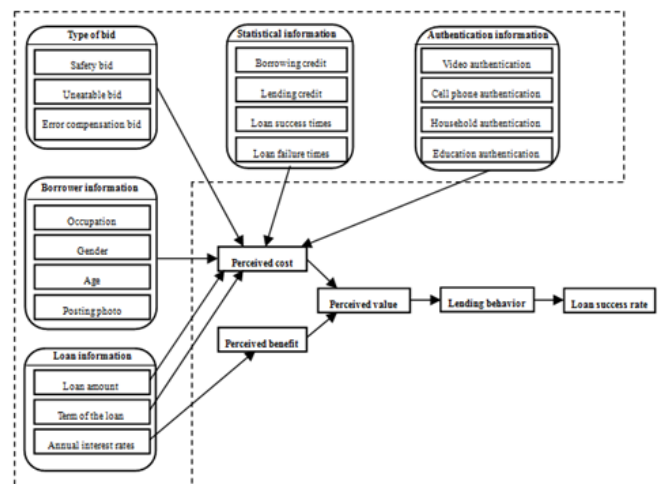


Figure 1: theoretical model diagram

3.2 Related assumptions about influencing factors of lending behavior

Seeing from figure 1, lenders through the analysis of the loan list information, form the perceived benefits and perceived costs of loan request, then compare the difference of them and determine their lending behavior, therefore, perceived benefits and perceived costs are the fundamental factors that affect the lender's lending behavior.

(a) The relationship between loan information and lender's perceived benefits

The higher the interest rate, the higher the lender's returns, and the higher the perceived benefits. So we make the following assumptions:

H1: the lender's perceived benefit is positively correlated with the annual interest rate.

(b) The relationship between loan information and lender's perceived cost

Usually the higher borrower's loan amount is, the greater repayment pressure in each period is, so the borrower will be more likely to default; in addition, according to historical experience, the default risk of long-term loan is higher than short-term loan. So we make the following assumptions:

H2: the lender's perceived cost is positively related to the loan amount and the term of the loan.

(c) The relationship between the personal information of the borrower and the lender's perceived cost

The gender, age, occupation and other factors will affect borrower's income level, thus affecting their repayment ability. Usually if borrowers' works are not stable, or borrowers are younger females, their repayment ability are weak. If the borrower does not post photos, at least it shows that borrower pays little attention to his loan, which will increase the corresponding default risk. In summary, the lender may believe that the loan default risk of the borrower

with above characteristics will be higher. So we make the following assumptions:

H3: lender's perceived cost will be higher, if borrower is a young female, doesn't have stable work, or doesn't post the photo.

(d) The relationship between the type of bid and the perceived cost of the lender

The loan risk of safety bid, uncashable bid and error compensation bid are low, because even if the borrower's repayment is overdue, or borrower defaults on the loan and so on, lender still can guarantee the safety of their income. These three types of bid's audits are more stringent, securities are higher, which will reduce the lender's loan risk. So we make the following assumptions:

H4: safety bid, uncashable bid and error compensation bid will reduce the lender's perceived cost.

(e) The relationship between the borrower's authentication information and the lender's perceived cost

The borrower's identity authentication can make lenders more clearly understand the borrower's identity information, because of fears that their identity information could be exposed, the borrower dare not readily default on the load, so the loan risk of the borrower with certain authentication information is low. So we make the following assumptions:

H5: lender's perceived cost will be lower, if the borrower has the video, cell phone, household or education authentication.

(f) The relationship between the historical statistical information of borrower's loan and the lender's perceived cost

In order to make the lender more aware of the borrower, PPDAl will record the historical statistical information of borrower's loan. In the historical information, the score of lending and borrowing credit are calculated according to the level of detail in the borrower's personal information and historical statistical information of borrower's lending and repayment, and the higher the score, the better credit, the lower the default risk. Moreover, to some extent, more loan

success times and fewer loan failure times show that the former lender's evaluation on borrower is better, and borrower's loan risk is relatively low. So we make the following assumptions:

H6: the lender's perceived cost is lower, if the score of borrower's lending credit or borrowing credit are higher, or borrower has more loan success times or fewer loan failure times.

4. Empirical Analysis

In this paper, we take the registered users of PPDAlP2P network lending platform as the research object, the research data also comes from this platform, which is mainly based on the following considerations:

(a) This platform started early, the transaction volume is large, which is helpful for data mining and model building. P2P network lending platform will generally have more loans which are withdrawn, in order to ensure the sample size, should choose the platform from which more data can be collected.

(b) PPDAl takes independent transaction as the as the main market transaction mode, which is more conducive to the study of the objective laws of P2P network lending.

(c) PPDAl's service objects are not limited to a specific area, covering the small and micro business owners and ordinary individuals in China's most provinces and cities, so that the samples will be more comprehensive.

4.1 Data preparation

4.1.1 Introduction to influencing variables

The main reference of the lender's investment decision is the borrower's loan list information, so the relevant variables in the borrower's loan list and the borrower's home page will be included in the analysis. Through the analysis of the interaction among the variables, as well as the impacts of variables on the lending behavior, the information of relevant variables obtained in this paper is shown in table 1.

Table 1: Information introduction about influencing variables

| Primary index | Secondary index | Interpretation of secondary index |
|-------------------------|--------------------|---|
| Statistical information | Lending credit | Lending credit scoring standards: as a lender, each bid plus two points, each interest received plus two points, receiving the full principal and interest plus two points, each overdue repayment minus 10 points. |
| | Borrowing credit | Borrowing credit scoring standards: ranking the borrowing credit according to the borrower's personal information, authentication information, repayment information and other information. 150 is the highest Borrowed credit score, from high to low is A, B, C, D, E, HR level, each increase of 25 points will increase the level of 1. |
| | Loan success times | Borrower's succeeding numbers of loan. |
| | Loan failure times | Borrower's failing numbers of loan. Failure reason: first, bid is abortive; second, although |

| | | |
|----------------------------|--------------------------|---|
| | | accomplishing the full scale of the loan amount, the loan request doesn't pass the scrutiny. |
| Authentication information | Video authentication | Video authentication requires the borrower to upload his video on the video authentication page. |
| | Household authentication | Ensure the authenticity of identity information. |
| | Education authentication | Education authentication need to authenticate the user's education. To ensure the authenticity of education information. |
| | Cellphone authentication | Cellphone real name authentication needs to authenticate the user's mobile phone number, to determine the normal use of mobile phones, and there are enough contacts. |
| Loan information | Loan amount | The loan amount is 1000-500000 Yuan. |
| | Annual interest rate | The scope of the annual interest rate is 5%-26.24% (4 times the base annual interest rate in 2011). |
| | Term of loan | The scope of the term of loan is 1-12 months. |
| Type of bid information | Safety bid | The borrower's lending credit score is above 200. |
| | Uncashable bid | Borrowed money can't cash, can only be used for the purpose of repayment and bid on PPDAl platform. |
| | Error compensation bid | As long as lenders find any errors or omissions in the process of loan audit, PPDAl will make up for all loss of users. |
| Borrower information | Posting photo | In the loan list, some borrowers will post their photos to increase credibility. |
| | Occupation | Occupation is divided into working class, online store shopkeeper, private owners, students, etc.. |
| | Gender | The borrower's gender is male or female. |
| | Age | Borrower's age is divided into 4 sections: 20-25 years old, 26-31 years old, 32-38 years old, older than the age of 39. |
| Loan state information | Fully paying off | The loan is successful and is fully paid off |
| | Success | Full Scale, and the loan has passed the scrutiny. |
| | Approval failure | Full Scale, but the loan doesn't pass the scrutiny. |
| | Abortive | Failing to accomplish the full scale of the loan amount in 7 days. |
| | withdrawn | For some reason, the borrower withdraws the loan applications. |

This paper is a study of lender's lending behavior, for the same loan application, within the specified time, the more the persons deciding to bid, the higher the possibility of Full Scale, and vice versa, so whether the full scale of the loan.

amount can be accomplished can be used as an index to measure the lender's lending behavior, so this paper selects whether the full scale of the loan amount is accomplished as the dependent variable, and takes the influencing variables in Table 1 as the independent variables, as shown in table 2.

Table 2: related variables

| Primary index | | Secondary index | | | |
|----------------------|----------------------------|---|--------------------------------------|--------------------------------------|--------------------------------------|
| Dependent variable | | whether the full scale of the loan amount is accomplished (y) | | | |
| Independent variable | Loan information | Annual interest rate (x_1) | Loan amount (x_2) | Term of loan (x_3) | |
| | Type of bid | Safety bid(x_4) | Uncashable bid(x_5) | Error compensation bid(x_6) | |
| | Borrower information | Posting photo (x_7) | Occupation(x_8) | Age(x_9) | Gender(x_{10}) |
| | Statistical information | Borrowing credit(x_{11}) | Lending credit(x_{12}) | Loan success times(x_{13}) | Loan failure times(x_{14}) |
| | Authentication information | Video authentication(x_{15}) | Household authentication(x_{16}) | Education authentication(x_{17}) | Cellphone authentication(x_{18}) |

4.1.2 Data extraction process

All loan information have been stored on a corresponding web page of PPDAl P2P network platform, and the end of this page's URL (Universal Resource Locator) is the number of the corresponding loan, such as <http://www.ppdai.com/list/602221>.

We use web crawler to grab the loan information on these pages, the principle is turning HTML page text into Document Object Model (DOM) for analysis, to obtain the required information.

Using this program, we grab 40757 pieces of loan data from PPDAl's website, the time span is July 22, 2014 to July 22, 2015, there are 3083 pieces of loan data of which loan state is withdrawn, because the reasons for the withdrawal cannot be determined, at the time of research will generate higher uncertainty, so we will remove these data in the process of logistic regression analysis, only 37674 pieces of loan data remaining.

4.2 Empirical design and related principles

4.2.1 Choice of regression model

When the dependent variable is a nominal variable, the

ordinary linear regression analysis is unable to deal with this situation, and the Logistic regression is used to deal with the case when the dependent variable is a nominal variable. In this paper, we choose the "Full Scale" as the dependent variable, and Gilkeson & Reynolds (2004) use Logistic regression to analyze how success rate of startup affects the success rate of auction^[15]. So this paper uses Logistic regression analysis. The Logistic regression analysis is described as follow:

$$P = P(Y = 1) = \frac{1}{1+e^{-z}} \quad (1)$$

Where:

P—Indicates the possibility of Full Scale under a given set of independent variables

z—Independent variable

The independent variable z is defined as follows:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (2)$$

After the logarithmic transformation, the above formula is:

$$\ln \frac{p(Y=1)}{p(Y=0)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (3)$$

Where:

x_1, x_2, \dots, x_k —Each independent variable in the Formula (2) and (3)

$\beta_1, \beta_2, \dots, \beta_k$ —The coefficients corresponding to each variable

β_0 —Constant term

ε —Stochastic disturbed term

4.2.2 Logarithm likelihood test

The log likelihood (LL) is the natural logarithm of the probability value, because the range of probability value is [0, 1], so the range of log likelihood is [0, -∞]. -2 log likelihood reflects the error caused by the factors outside of independent variables in the model, because it approximates to Chi-square distribution, so we can use this feature to deal with the problem that whether the change of dependent variable caused by the factors outside of independent variables in the model is significant, so it is also called Badness-of-fit Chi-square. When the value of -2 log likelihood is greater than a certain value, the change of the dependent variable can't be explained by the independent variable. -2 log likelihood is calculated as follow:

$$-2LL = -2 \sum_{i=1}^n [y_i \ln(\hat{p}_i) + (1 - y_i) \ln(1 - \hat{p}_i)] \quad (4)$$

4.2.3 Two-step clustering

The two-step clustering method is a kind of intelligent clustering method, and it arises with the development of artificial intelligence, it is characterized by: 1. It can handle both continuous and nominal variables; 2. It can handle large

sample data and has a high operation efficiency; 3. It can automatically select the number of clusters according to the Akaike information criterion (AIC) or Bayesian information criterion (BIC). Combined with the characteristics of the data in this paper, we choose two-step clustering method. In the two-step clustering method, the distance between cluster i and cluster j is calculated as follows:

$$d(i, j) = \varepsilon_i + \varepsilon_j - \varepsilon_{(i,j)} \quad (5)$$

Where :

$$\varepsilon_v = -N_v \left[\sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2) + \sum_{k=1}^{K^B} E_{vkl} \right] \quad (6)$$

$$E_{vkl} = - \sum_{l=1}^{L_k} \frac{N_{vkl}}{N_v} \log \frac{N_{vkl}}{N_v} \quad (7)$$

$\varepsilon_{(i,j)}$ —The distance between cluster i and cluster j

K^A —The number of continuous variables

K^B —The number of nominal variables

$\hat{\sigma}_k^2$ —The estimation variance of the kth continuous variable

$\hat{\sigma}_{vk}^2$ —The estimation variance of the kth continuous variable in cluster v

L_k —The number of classification of the kth nominal variable

N_v —The sample size of cluster v

N_{vkl} —The sample size of the kth nominal variable whose value is l in cluster v

4.2.4 Determination of cluster number

Two-step clustering method is generally used AIC or BIC these two indicators to calculate whether the cluster number is suitable. The smaller the two indicators, the better the clustering results. AIC and BIC are calculated as follows:

$$AIC(J) = -2 \sum_{j=1}^J \varepsilon_j + 2m_j \quad (8)$$

$$BIC(J) = -2 \sum_{j=1}^J \varepsilon_j + 2m_j \log(N) \quad (9)$$

Where:

$$m_j = J[2K^A + \sum_{k=1}^{K^B} (L_k - 1)] \quad (10)$$

J—Cluster number

N—Sample size

$\sum_{j=1}^J \varepsilon_j$ —The maximum value of the likelihood function

Because AIC will cause over-fitting phenomena when there are too many parameters in the model, resulting in lack of rigor, so we choose BIC to determine the cluster number.

4.3 Logistic regression analysis

4.3.1 Coding of nominal variables and basic statistical characteristics of variables

The coding information of nominal variables and the descriptive statistics of continuous variables are shown in table 3 and 4, and the loan state information is shown in table 5.

Table 3: Coding information of nominal variables

| Nominal variable | classification | Number | Percentage (%) | Code | | | |
|--------------------------|-------------------------|--------|----------------|------|-----|-----|-----|
| | | | | (1) | (2) | (3) | (4) |
| Occupation | Working class | 20379 | 54.09 | 1 | 0 | 0 | 0 |
| | Others | 4608 | 12.23 | 0 | 1 | 0 | 0 |
| | Private owners | 8056 | 21.38 | 0 | 0 | 1 | 0 |
| | Online store Shopkeeper | 2886 | 7.66 | 0 | 0 | 0 | 1 |
| | Student | 1745 | 4.63 | 0 | 0 | 0 | 0 |
| Age | 20-25 | 11691 | 31.03 | 1 | 0 | 0 | |
| | 26-31 | 14401 | 38.23 | 0 | 1 | 0 | |
| | 32-38 | 8155 | 21.65 | 0 | 0 | 1 | |
| | >39 | 3427 | 9.10 | 0 | 0 | 0 | |
| Error compensation bid | No | 28805 | 76.46 | 0 | | | |
| | Yes | 8869 | 23.54 | 1 | | | |
| Safety bid | No | 36551 | 97.10 | 0 | | | |
| | Yes | 1123 | 2.90 | 1 | | | |
| Uncashable bid | No | 36439 | 96.72 | 0 | | | |
| | Yes | 1235 | 3.28 | 1 | | | |
| Gender | Male | 31910 | 84.70 | 1 | | | |
| | Female | 5764 | 15.30 | 0 | | | |
| Video authentication | No | 27715 | 73.57 | 0 | | | |
| | Yes | 9959 | 26.43 | 1 | | | |
| Cellphone authentication | No | 27011 | 71.70 | 0 | | | |
| | Yes | 10663 | 28.30 | 1 | | | |
| Household authentication | No | 25876 | 68.68 | 0 | | | |
| | Yes | 11798 | 31.32 | 1 | | | |
| Education authentication | No | 33835 | 89.81 | 0 | | | |
| | Yes | 3839 | 10.19 | 1 | | | |
| Posting photo | No | 18308 | 48.60 | 0 | | | |
| | Yes | 19366 | 51.40 | 1 | | | |
| Full Scale | No | 16831 | 44.68 | 0 | | | |
| | Yes | 20843 | 55.32 | 1 | | | |

Table 4: Descriptive statistics of continuous variables

| Continuous variables | Range | Min | Max | Mean | S.D. | Variance |
|----------------------|--------|-------|--------|---------|----------|--------------|
| Loan amount | 399000 | 1000 | 400000 | 5967.20 | 18956.59 | 359352437.59 |
| Annual interest rate | 16.00% | 8.00% | 24.00% | 19.22% | 3.30% | 10.90 |
| Term of loan | 11 | 1 | 12 | 8.00 | 3.22 | 10.37 |
| Lending credit | 96838 | 0 | 96838 | 316.43 | 2245.67 | 5043042.32 |
| Borrowing credit | 140 | 10 | 150 | 26.29 | 19.52 | 381.14 |
| Loan success times | 110 | 0 | 110 | 1.72 | 3.86 | 14.88 |
| Loan failure times | 21 | 0 | 21 | 1.71 | 1.80 | 3.23 |

Table 5: Loan state information

| | Success | Abortive | Approval failure | Withdrawn | Fully paying off | Total |
|----------------|---------|----------|------------------|-----------|------------------|--------|
| Number | 6822 | 16831 | 8757 | 3083 | 5264 | 40757 |
| Percentage (%) | 16.74 | 41.30 | 21.49 | 7.56 | 12.91 | 100.00 |

According to the information provided by table 3, 4 and 5, we can see that the average loan amount is 5 967.20 Yuan, the loan amount is relatively small, the average term of loan is 8 months, a shorter period of time and these two characteristics are in line with the P2P network lending's market positioning—small short-term lending.

But the average borrowing rate is 19.22%, is 3.2 times the 2014 China's financial institutions annual lending rate—6%, so the borrowing costs is relatively high;

Seeing from the borrower information, man accounts for 84.70% of the borrowers, age 38 and below accounts for 90.90%, working class accounts for 54.09%, the private

owner accounted for 21.38%, indicating that the main traders on PPDAl platform are young male workers and private owners.

Seeing from the type of bid, 70.28% of the borrowers select only ordinary standard. As for special type of bid, most people will choose the error compensation bid, accounting for 23.54%, safety bid and uncashable bid accounted for only 2.90% and 3.28% respectively.

Seeing from the historical statistical information, average lending credit score is 316.43, but the standard deviation is as high as 2245.67, indicating the lending credit disparities between borrowers. The average borrowing credit is 26.29, which is just above HR level, which is the minimum credit rating. The loan success times are roughly equivalent to the loan failure times; their mean values are 1.72 and 1.71 respectively.

Seeing from the authentication information, borrowers rarely choose education authentication, accounts for only 10.19%, while the number of choosing video, cellphone and household authentication account for 26.43%, 28.30% and 31.32% respectively. Based on the type of bid information, historical statistical information and authentication information, we find that the majority of borrowers' credit

rating is low, and most borrowers are unwilling to choose the special type of bid, only choosing less identity authentication method.

Seeing from the loan state Information, the Full Scale accounts for 51.14%, it is the sum of "fully paying off", "success" and "approval failure" these three loan states, compared to "abortive" loan state, which accounts for 41.3%, indicating that the lender's lending behavior is still active. It is important to note that the loan success is only the sum of "fully paying off" and "success" these two loan states, accounting for only 29.65%, because "approval failure" accounts for 21.49%, the higher rate of Full Scale does not translate into the higher corresponding loan success rate, resulting in the low efficiency of market transactions, and money and time are greatly wasted, this problem should cause the attention of traders and platform.

4.3.2 Related test of analysis correlation

We take whether the full scale of the loan amount is accomplished as the dependent variable, other variables as independent variables, using logistic regression to test the influencing factors of lender's lending behavior on China's P2P network lending platform by applying SPSS20.0, the results are as shown in table 6.

Table 6: Comprehensive test of all variables' coefficients and the fitness of model

| χ^2 | df | -2log likelihood | Nagelkerke R ² | Sig. |
|-----------|----|------------------|---------------------------|-------|
| 43435.089 | 23 | 8364.106a | 0.916 | 0.000 |

Seeing from table 6, the value of Sig. is 0, far less than 0.05, so we should reject the null hypothesis, namely, the regression coefficients of all variables are not 0 at the same time, so there is a significant linear relationship between all independent variables as a whole and the dependent variable, namely, the model is effective as a whole. When the sample

size is 37674, the corresponding value of -2log likelihood is 8364.106, and the value of Nagelkerke R² is 0.916, which indicates that the 91.6% of the change of dependent variable can be explained by independent variables, so this model has a higher fitting degree.

Table 7: Hosmer & Lemeshow test results

| Group | Abortive | | Full Scale | | Total observed |
|-------|----------|----------|------------|----------|----------------|
| | Observed | Expected | Observed | Expected | |
| 1 | 3774 | 3770.398 | 4 | 7.602 | 3778 |
| 2 | 3759 | 3749.504 | 8 | 17.496 | 3767 |
| 3 | 3753 | 3712.999 | 18 | 58.001 | 3771 |
| 4 | 3566 | 3536.445 | 201 | 230.555 | 3767 |
| 5 | 1756 | 1885.487 | 2011 | 1881.513 | 3767 |
| 6 | 192 | 159.515 | 3575 | 3607.485 | 3767 |
| 7 | 21 | 14.447 | 3746 | 3752.553 | 3767 |
| 8 | 6 | 1.979 | 3761 | 3765.021 | 3767 |
| 9 | 2 | 0.217 | 3764 | 3765.783 | 3766 |
| 10 | 2 | 0.008 | 3755 | 3756.992 | 3757 |

According to the results of Hosmer & Lemeshow test, we find that the expected value of the Abortive and the Full Scale is broadly consistent with the observed value, so the fitting degree of the model is good.

Table 8: Classification results of the logistic regression

| Observed | | Expected | | |
|----------------------|-------|----------|------------|----------------|
| | | Abortive | Full Scale | Percentage (%) |
| Abortive | 16831 | 16075 | 756 | 95.5 |
| Full Scale | 20843 | 926 | 19917 | 95.6 |
| Total percentage (%) | | | | 95.5 |

According to table 8, the observed numbers of the abortive loan applications are 16831, and corresponding expected numbers are 16075, so the prediction accuracy rate is 95.5%; the observed numbers of the Full Scale loan applications are 20843, and corresponding expected numbers are 19917, so the prediction accuracy rate is 95.6%. Therefore, the application of the model in the “abortive” and “Full Scale” loan request forecasting obtains better forecasting results.

4.3.3 Determination of regression model

According to table 9, the sig. of all variables all are less than 0.05, so all of the independent variables should be retained in the model. We input the values of each independent variable’s coefficient and the constant term into the formula (1), so we obtain the calculation formula of the possibility of Full Scale:

$$P = 1/(1 + e^{(-(-8.982 + 0.319x_1 - 0.064x_3 + 2.35x_4 + 2.325x_5 + 11.046x_6 + 2.279x_7 - 1.970x_{8(1)} - 2.088x_{8(2)} - 1.896x_{8(3)} - 0.666x_{8(4)} - 0.386x_{9(1)} - 0.248x_{9(2)} - 0.272x_{9(3)} + 2.279x_{10} + 0.117x_{11} + 0.004x_{12} + 0.609x_{13} - 0.376x_{14} - 0.371x_{15} + 0.787x_{16} + 0.384x_{17} - 0.984x_{18}))}) \quad (11)$$

Where:

$x_{8(1)}, x_{8(2)}, x_{8(3)}, x_{8(4)}$ —Representing working class, others, private owner and online store shopkeeper respectively
 $x_{9(1)}, x_{9(2)}, x_{9(3)}$ —Representing 20-25 age group, 26-31 age group and 32-38 age group respectively

Table 9: Regression results of all independent variables

| Variable | B | S.E | Wals | df | Sig. | Exp(B) |
|-----------------------------|--------|-------|---------|----|-------|--------|
| Loan amount | 0.000 | 0.000 | 22.186 | 1 | 0.000 | 1.000 |
| Annual interest rate | 0.319 | 0.014 | 527.164 | 1 | 0.000 | 1.376 |
| Term of loan | -0.064 | 0.009 | 48.458 | 1 | 0.000 | 0.938 |
| Lending credit | 0.004 | 0.000 | 94.687 | 1 | 0.000 | 1.004 |
| Borrowing credit | 0.117 | 0.060 | 408.150 | 1 | 0.000 | 1.124 |
| Loan success times | 0.609 | 0.030 | 400.269 | 1 | 0.000 | 1.838 |
| Loan failure times | -0.376 | 0.017 | 490.668 | 1 | 0.000 | 0.687 |
| Posting photo(1) | 2.279 | 0.081 | 798.394 | 1 | 0.000 | 9.765 |
| Video authentication(1) | -0.371 | 0.095 | 15.101 | 1 | 0.000 | 0.690 |
| Cellphone authentication(1) | -0.984 | 0.097 | 102.860 | 1 | 0.000 | 0.374 |
| Household authentication(1) | 0.787 | 0.072 | 120.080 | 1 | 0.000 | 2.196 |
| Education authentication(1) | 0.384 | 0.111 | 11.902 | 1 | 0.001 | 1.496 |
| Gender(1) | -0.320 | 0.077 | 17.193 | 1 | 0.000 | 0.726 |
| Age | | | 10.517 | 3 | 0.015 | |
| 20-25(1) | -0.386 | 0.125 | 9.574 | 1 | 0.002 | 0.680 |
| 26-31(2) | -0.248 | 0.122 | 4.098 | 1 | 0.043 | 0.781 |
| 32-38(3) | -0.272 | 0.131 | 4.298 | 1 | 0.038 | 0.761 |
| Occupation | | | 453.542 | 4 | 0.000 | |
| Working class(1) | -1.970 | 0.109 | 328.041 | 1 | 0.000 | 0.140 |
| Others(2) | -2.088 | 0.134 | 240.993 | 1 | 0.000 | 0.124 |
| Private owner(3) | -1.896 | 0.127 | 221.946 | 1 | 0.000 | 0.150 |
| Online store shopkeeper (4) | -0.666 | 0.137 | 23.725 | 1 | 0.000 | 0.514 |
| Uncashable bid(1) | 2.325 | 0.504 | 21.271 | 1 | 0.000 | 10.225 |
| Safety bid(1) | 2.350 | 1.188 | 3.911 | 1 | 0.048 | 10.484 |
| Error compensation bid(1) | 11.046 | 0.390 | 803.317 | 1 | 0.000 | 62.709 |
| Constant term | -8.982 | 0.366 | 602.570 | 1 | 0.000 | 0.000 |

4.3.4 Results analysis

According to the value of coefficient B and Exp (B) of each independent variable in table 9, to analyze the effect of each variable on the Full Scale:

(a) Variables about the loan information

Annual interest rate has a positive impact on the Full Scale. The higher the interest rate, the higher the lender's income accordingly, that is, the higher the lender's perceived benefit, and the annual interest rate increases by 1%, the possibility of

Full Scale increases by 37.6%, showing that the annual interest rate has a strong impact on the Full Scale. From the above, the results support the hypothesis H1: the lender’s perceived benefit is positively correlated with the annual interest rate.

The loan amount basically has no influence on the Full Scale; this seems to be inconsistent with our common sense. Usually the higher the loan amount, the lower the possibility of obtaining a loan, but PPDAl completely breaks this cognition, because the pattern of lending in PPDAl is similar to crowd-funding, whatever the loan amount is, each lender

can choose their own investment amount according to their cognition on the loan list information and their economic strength, so that the lender will ignore the loan amount requested in the loan list, so the loan amount has very little influence on the Full Scale.

The impact of the term of loan on the Full Scale is negative. The longer the duration of the loan, the higher the lenders risk, for each additional month, the possibility of the Full Scale decreases by 6.2%, indicating that the lender's perceived cost of the term of loan period is relatively low.

From the above, the results partially support the hypothesis H2: the lender's perceived cost is positively related to the loan amount and the term of the loan. Long term of loan will increase the lender's perceived cost, but the loan amount has no effect on it.

(b) Variables about the borrower information

The effect of gender on the Full Scale is negative. Women are 27.4% more likely than men to be Full Scale. This suggests that people are more likely to lend money to women, and it may be thought that women have a lower risk of default than men, which reduces the perceived cost of borrowers.

The effect of age on the Full Scale is positive. According to table 4-9, the regression coefficients in the 20-25 age group is -0.386, 26-31 age group is -0.248, 32-38 age group is -0.272, older than 39 age group is 1, and generally show the trend that these coefficients increase with the age increasing, so we say that the effect of age on the Full Scale is positive. Compared to older than 39 age group, the possibility of the Full Scale in 20-25 age group decreases by 32%, 26-31 age group decreases by 21.9%, 32-38 age group decreases by 23.9%, indicating that lenders are more willing to invest in older borrowers, because their average income level and repayment ability are higher, so older borrowers are more favored by lenders.

The effect of occupation on the Full Scale is negative. To some extent, the type of occupation will reflect the economic strength of borrowers, because of their lower risk of default, people are more willing to borrow money to the people with higher economic strength. According to table 4-9, the regression coefficient of the working class is -1.97, the private owner is -1.896, the online store shopkeeper is -0.666, the "others" is -2.088 and the student is 1. Seeing from the regression coefficient, it is clear that the regression coefficient of the "others" is the smallest, without indicating the occupation information will significantly increase the lender's skepticism about the repayment ability of the borrower, thereby increasing the lender's perceived cost, so the negative effect of the "others" on the Full Scale is the biggest. A surprising finding is that comparing to the student borrower, the working class's possibility of the Full Scale will decrease by 86%, the private owner will decrease by 85%, the online store shopkeeper will decrease by 48.6%, the "others" will decrease by 87.6%, which indicates that lenders are more willing to invest in students. Although students

have no fixed source of income, but they have parents as natural guarantor, and they have to face the pressure of enrollment and finding employment, so the negative effects caused by the loan default will cause their higher default cost, thereby the loan risk of the student is relatively low. In recent years, there have been a number of loan platforms catering specifically for students in China, such as Fenqile, Ufenqi, Xiaoyuan52, etc., which also confirms the above explanation to some extent.

Posting photo has a positive impact on the Full Scale. The possibility of the Full Scale of the person who has posted the photo is 9.765 times the possibility of the person without posting his photo, so the impact of posting photo is very large. Because PPDAl after all is a virtual network loan platform, both lenders and borrowers can't meet face to face and get to know each other, so on the P2P network lending platform, lenders' perceived cost is higher the cost in the realistic environment. Posting borrowers' photo can give lenders a sense of security and credibility, thereby reducing borrowers' perceived cost.

The results partially support the hypothesis H3: lender's perceived cost will be higher, if borrower is a young female, doesn't have stable work, or doesn't post the photo. The lender has a greater perceived cost for the 20-25 age group, the occupation of the "others", and the person who does not post the photo, which supports the hypothesis H3, but the lender is more inclined to invest in women, which is inconsistent with the original hypothesis.

(c) Variables about the type of bid information

The uncashable bid, safety bid and error compensation bid all have positive impacts on the Full Scale. Although choosing these special types of bid will pay a fee to PPDAl, but from the view of the result, all these special types of bid have significant positive impacts on the Full Scale, the uncashable bid, safety bid and error compensation bid can increase the possibility of the Full Scale by 10.225 times, 10.484 times and 62.709 times respectively. According to the type of bid information in table 4-1, all these special types of bid have the corresponding risk prevention mechanisms, their safeties are the highest on PPDAl platform, thus reducing the perceived cost of the lender.

The results support the original hypothesis H4: safety bid, uncashable bid and error compensation bid will reduce the lender's perceived cost.

(d) Variables about the authentication information

The impacts of the household authentication and education authentication on the Full Scale are both positive, the impacts of cellphone authentication and video authentication on the Full Scale are both negative. On PPDAl platform, the borrower must pay a certain fee for each type of authentication, and prepare some audit materials, so some borrowers will not choose to authenticate their information. PPDAl promotes users' information authentication by saying

that authentication will enhance users' credibility and awareness, thereby increasing their success rate of loan. Do all the types of authentication have positive effects on the Full Scale? According to the results of the data analysis, we find that video authentication and mobile phone authentication have negative effects on the Full Scale, indicating that the lender is not very concerned about these two authentications. The household authentication and education authentication will increase the possibility of the Full Scale by 2.196 times and 1.49 times respectively, compared with the video authentication and mobile phone authentication; the lender will pay more attention to the authenticity of borrower's household and education information.

The results partially support the hypothesis H5: lender's perceived cost will be lower, if the borrower has the video, cellphone, and household or education authentication. The household authentication and education authentication can reduce people's perceived cost and support the original hypothesis, but video authentication and mobile phone authentication can't reduce the perceived cost of the lender.

(e) Variables about the historical statistical information

The lending credit has a faint positive impact on the Full Scale, for each additional point of the lending credit; the likelihood of the Full Scale will only increase by 0.4%. The lending credit reflects the historical lending situations of the borrower; the impact of this variable is faint, indicating that lenders do not care too much about the borrower's lending situation.

The borrowing credit has a positive impact on the Full Scale; the likelihood of the Full Scale will increase by 12.4% for each additional point of borrowing credit. The borrowing credit reflects the borrower's historical borrowing situations; this variable is more relevant to the borrower's repayment ability, so compared with the lending credit, it has a more significant impact on the Full Scale.

The loan success times have a positive impact on the Full Scale, it also reflects the borrower's borrowing history, and more intuitively, so the impact is relatively large, for each additional time of the loan success times, the likelihood of the Full Scale will increase by 83.8%. The more the borrower's loan success times, the more the lender's sense of

security, so a relatively high level of the loan success times will reduce the perceived cost of the lender, thereby stimulating the lending behavior.

The loan failure times have a negative impact on the Full Scale, and for each additional time of the loan failure times, the likelihood of the Full Scale will decrease by 31.3%. Because the more the borrower's loan failure times, the lower the lender's investment willingness in this borrower, so this variable will directly affect the lender's lending behavior. Through the analysis the impacts of the loan success / failure times on the Full Scale, we find that there exists the herding behavior in China's P2P network lending, namely the more the loan success time of a person, the higher the success rate of his later loan and vice versa, which is consistent with Tao Wen & Long Jiangcheng's (2015) research results^[16]. Compared with the loan failure times, the loan success times have a more significant impact on the Full Scale, indicating that lenders tend to pay more attention to borrower's succeeding numbers of loan.

The results support the hypothesis H6: the lender's perceived cost is lower, if the score of borrower's lending credit or borrowing credit are higher, or borrower has more loan success times or fewer loan failure times.

4.4 Two-step cluster analysis

On the PPDPAIP2P network lending platform, there are a large number of users with a variety of loan activities every day. Every loan list released contains borrower's personal information, historical statistical information, loan information and other important information, through the two-step cluster analysis of these information, we can classify borrowers, research the characteristic difference between the different categories of borrowers, hope to find out whether there are differences in the Full Scale and other variables between the different categories of borrowers.

We use SPSS20.0 to conduct two-step cluster analysis on a total of 40757 pieces of data about 18 variables, and use BIC to select the cluster number automatically.

In this paper, we choose the loan state as the evaluation field in cluster analysis, and the distribution information of the loan state in each cluster is shown in table 10:

Table 10: Distribution information of the loan state in each cluster

| | | Success | Abortive | Approval failure | Withdrawn | Full paying off | Total |
|-----------|----------------|---------|----------|------------------|-----------|-----------------|-------|
| Cluster 1 | Number | 5575 | 3038 | 5499 | 2266 | 4525 | 20903 |
| | Percentage (%) | 26.67 | 14.53 | 26.31 | 10.84 | 21.65 | 100 |
| Cluster 2 | Number | 1247 | 13793 | 3258 | 817 | 739 | 19854 |
| | Percentage (%) | 6.28 | 69.47 | 16.41 | 4.12 | 3.72 | 100 |

According to the distribution information of each loan state in the table 10, we can find that in cluster 1, success, approval failure and fully paying off these three Full Scale states are the majority, the Full Scale rate (74.63%) is significantly more than the abortive rate (14.53%), and obviously more

than the Full Scale rate (26.41%) in the cluster 2; In the cluster 2, the abortive state is the main part, accounting for 69.47%, and other states are the minority. Therefore, the clustering results will be automatically divided into the cluster 1 in which the Full Scale states are the majority and

the cluster 2 in which the abortive state is the majority, and the cluster results show that there is good consistency for the internal information in each cluster.

Summary information of the independent variables in two clusters is shown in table 11:

Table 11: Clustering information

| | | | | | |
|------------|------------------------|--------------------------|----------------------|--------------------------|--------------------------|
| | Lending credit | Loan failure times | Posting photo | Occupation | Loan success times |
| Importance | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Cluster 1 | Mean : 39.37 | Mean : 2.22 | Yes : 96.8% | Working class : 58.5% | Mean : 3.20 |
| Cluster 2 | Mean : 12.53 | Mean : 1.17 | No : 93.6% | Working class : 49.8% | Mean : 0.17 |
| | Term of loan | Household authentication | Video authentication | Cellphone authentication | Education authentication |
| Importance | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Cluster 1 | Mean : 7.26 | Yes : 60.2% | Yes : 50.9% | Yes : 52.8% | No : 80.0% |
| Cluster 2 | Mean : 8.77 | No : 98.7% | No : 98.6% | No : 97.0% | No : 99.4% |
| | Uncashable bid | Safety bid | Loan amount | Lending credit | Age |
| Importance | 0.92 | 0.84 | 0.69 | 0.51 | 0.27 |
| Cluster 1 | No : 93.6% | No : 94.2% | Mean : 8792 | Mean : 603.00 | 26-31:40% |
| Cluster 2 | No : 100% | No : 100% | Mean : 2991 | Mean : 13.79 | 26-31:36% |
| | Annul interesting rate | Error compensation bid | | Gender | |
| Importance | 0.07 | 0.05 | | 0.49e-4 | |
| Cluster 1 | Mean : 19.38 | No : 79.9% | | Male : 84.6% | |
| Cluster 2 | Mean : 19.05 | No : 76.5% | | Male : 84.7% | |

According to the table 11, we sort the independent variables by the importance influencing the clustering results, and the most important variables are the borrowing credit, loan failure times, posting photo, occupation, loan success times, term of loan, household authentication, video authentication, cellphone authentication and education authentication, their importance values are all 1; the least important variable is gender, its importance value is almost 0; the importance of the rest variables are in an intermediate position broadly.

In terms of the historical statistical information, the mean value of the borrowing credit in cluster 1 is 39.37, so its credit rank is E, and the mean value of the borrowing credit in cluster 2 is 12.53, its credit rank is HR (the lowest rank), that is, the credit rank in cluster 1 is higher than the credit rank in cluster 2. The importance of the lending credit is relatively low, its value is 0.51, but there is a large gap of the lending credit score between the cluster 1 and 2, the lending credit score in cluster 1 is 603, and corresponding score in cluster 2 is only 13.79, so that the person with the higher borrowing credit may also has the higher lending credit. The mean value of the loan failure times in cluster 1 is 2.22, it is bigger than cluster 2 (1.17). However, the mean value of the loan success times in cluster 1 is 3.2, and the corresponding value in cluster 2 is only 0.17, which shows that borrowers in cluster 1 will launch more borrowing behavior, so their loan success and failure times are both higher the cluster 2.

In terms of the authentication information, in cluster 1, the proportion of the household, video and cellphone authentications are all above 50%, although the proportion of education authentication is only 20%, but it accounts for almost all the samples of education authentication; in cluster 2, the proportion of these four authentications are all less than 3%.

In terms of the loan information, the terms of the both cluster are similar, cluster 1 is 7.26 months, and cluster 2 is 8.77 months. The importance of the loan amount is 0.69, it is not very high, and the mean value of the loan amount in cluster 1 is 8792, which is obviously higher than 2991—the mean value of the loan amount in cluster 2.

In terms of the type of bid information, clustering 1 occupies all the uncashable bid and safety bid, and there is no both uncashable bid and safety bid in cluster 2.

In terms of the borrower information, there is a significant difference in posting photo between cluster 1 and 2, 96.8% of borrowers in cluster 1 choose to post their photos, but this number in cluster 2 is only 6.4%.

The influences of the age, annual interest rate, error compensation bid and gender on the cluster results are all very small, and the differences of these four variables between the two clusters are all very small, too. According to the table 11, we find that there are significant differences in the data of some variables between the two clusters, and the differences between these variables may be the key factors that affect the Full Scale and the loan success.

In summary, we can draw the following conclusions: the quality of borrowers in cluster 1 is high, they will refine the information in their loan lists, try to make all kinds of authentication and more actively select the special type of bid, thereby obtaining the higher lending and borrowing credit and the more times of loan success; the quality of borrowers in cluster 2 is relatively low, most of them lack the will to improve the information in their loan lists, get less authentication and special type of bid than borrowers in cluster 1, so their credit rank and the loan success times are both low. To a certain extent, the hypotheses H4, H5 and H6 are proved by the above conclusions.

5. Conclusion

This paper first introduces the related theory and model about the influencing factors of lenders' lending behavior in China's P2P network lending, and puts forward the corresponding hypothesis; then we establish a logistic regression model based on the data from July 22, 2014 to July 22, 2015 on PPDAl, a P2P network lending platform, and use this model to conduct the empirical analysis on the influencing factors of lenders' lending behavior; finally, we use two-step cluster method to analyze the differences of all independent variables between the two clusters and get the following conclusion:

- 1) The loan success rate is relatively low on PPDAl platform, only 29.65%. This will result in the low efficiency of market transactions and cause a great waste of the capital and time of both sides of the transaction. This issue should be paid attention to by the P2P network lending platforms and traders.
- 2) Special type of bid, posting photo and the loan success/failure times all have important impacts on the success of the loan. All of these factors will reduce the perceived cost of the lender, and then improve the borrower's success rate of the loan; the loan amount basically has no influence on the success of the loan.
- 3) The borrowing credit, loan failure times, posting photo, occupation, loan success times, term of loan, household authentication, video authentication, cellphone authentication and education certification are the most important influencing factors on the cluster results, and there are significant differences in these variables between the two clusters of borrowers.

In order to improve the success rate of the loan, borrowers can draw lessons from the analysis of this article; improve the relevant factors in their loan lists, such as increasing the borrowing interest rates, posting their photos, authenticating their personal information and other measures, to improve the success rate of the loan. At the same time, for those who have a higher success rate of borrowing, can suitably reduce their borrowing interest rates to reduce their borrowing costs.

Acknowledgements

This research is supported by Central University Basic Research and Operating Expenses of Special Funding (NO. 72103779)

Reference

- [1] Iyer R., Khwaja A. I., Luttmer E. F. P., & Shue K.. "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending?" Cambridge: Harvard University, 2009.
- [2] Pope D. G., & Sydnor J. R.. "What's in a Picture? Evidence of Discrimination from Prosper. Com".

- Journal of Human Resources, 46(1), pp. 53-92, 2011.
- [3] Laura G., Yuliya K. L. "When Can a Photo Increase Credit? The Impact of Lender and Borrower Profiles on Online Peer-to-Peer Loans". Journal of Behavioral and Experimental Finance, 2014, 2(2), pp. 44-58, 2014.
- [4] Greiner M. E., & Wang H.. "The Role of Social Capital in People-to-People Lending Marketplaces". Phoenix: Association for Information Systems, 2009.
- [5] Collier B., & Hampshire R.. "Sending Mixed Signals: Multilevel Reputation Effects in Peer-to-Peer Lending Markets". Savannah: ACM Conference on Computer Supported Cooperative Work, 2010.
- [6] Klafft M.. "Peer to Peer Lending: Auctioning Microcredits over the Internet". Social Science Electronic Publishing, 2008.
- [7] Lu Xin, Li Minhui. "The Business Model and Risk Management in P2P network lending". Reform, (2), pp. 60-68, 2015.
- [8] Liu Hui, Shen Qingjie. "Research on the Risk and Regulation of P2P network lending in China". Research on Financial and Economic Issues, (1), pp. 52-59, 2015.
- [9] Wang Changjiang, Yang Jinye. "Research on the Risk and Supervision Mode of P2P network lending". Economic Review, (4), pp. 98-102, 2015.
- [10] Li Guangming, Zhou Huan. "An Empirical Study on the Borrower's Debt Characteristics in P2P network lending". Commercial Times, (1), pp. 41-42, 2015.
- [11] [11] Xin Xian. "An Exploration of P2P Operation Mode". Market Modernization Magazine, 2009, (19), pp. 19-22, 2009
- [12] [12] Peter, J. Paul, Sr. "Tarpey. A Comparative Analysis of Three Consumer Decision Strategies". Journal of Consumer Research, 2(1), pp. 29-37, 1975.
- [13] [13] S. Monroe, B. Kent, R. Krishnan. "The Effect of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value, Transaction Value, and Behavioral Intentions". Journal of Marketing, 62(2), pp. 46-59, 1985.
- [14] [14] C. Armitage, M. Conner. "Efficacy of the Theory of Planned Behavior: A Meta-analytic Review". British Journal of Social Psychology, (40), pp. 471-499, 2001.
- [15] [15] J.H. Gilkeson, K. Reynolds. "Determinants of Internet Auction Success and Closing Price: an Exploratory Study". Psychology & Marketing, 20(6), pp. 537-566, 2004.
- [16] [16] Tao Wen, Long Jiangcheng. "Herding Behavior in Peer to Peer Lending Market of China". International Journal of Science and Research, 6(2), pp. 1904-1910, 2015.

Corresponding Author Profile



Wang Kecheng (1982--), male, Shandong Qingdao, Xidian University, College of economics and management, master, research direction is technical economics and management.

Co Author Profile

Zhang Weihong (1966--), female, Beijing, Xidian University, College of economics and management, professor and master's tutor, research direction is corporate finance and project investment & financing management.

Wang Ximeng (1991--), female, Beijing, Xidian University, College of economics and management, master, research direction is the management of enterprise.

