Modelling Consumer Financing Behaviour in China

Lu Han

Modelling Consumer Financing Behaviour in China

Modelling Consumer Financing Behaviour in China

By

Lu Han

Cambridge
Scholars
Publishing

Modelling Consumer Financing Behaviour in China

By Lu Han

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-4371-4
ISBN (13): 978-1-5275-4371-3

TABLE OF CONTENTS

Acknowledgments

The work was supported by the Program for Innovation Research in the Central University of Finance and Economics and the Visiting Scholar Grant Program of the China Scholarship Council for Han (No. 201806495014).

My most sincere thanks to the Internet Finance Laboratory of the PBC School of Finance in Tsinghua University, the Credit Card Center in the Industrial and Commercial Bank of China, the Credit Ease Corporation and the Credit Reference Center in the People's Bank of China for their data support.

Finally, I would like to thank Helen Edwards at Cambridge Scholars Publishing for the editing support and thank Maya Berger for the proofreading. Anyway, if there is any mistake, it is my responsibility.

PREFACE

Breaking the bottle neck of small and medium-sized enterprise (SME) financing and promoting consumption by lifting consumer credit have become important ways to realize the transformation from whole sale finance to retail finance in China. With the development of the finance industry, competition and profit growth have shifted to consumer finance. Consumer finance studies consumer preferences and behaviors to discuss financial markets, financial products and financial policies.

According to the previous research, consumer finance is more unique than corporate finance. Firstly, data on corporate finance can be obtained through financial statements and reports, while data on consumer finance often need to be investigated. Secondly, the main subject of corporate finance is the company, which is consistent with the assumption of maximizing financial interests; while the main subject of consumer finance is human beings. The attitude and behavior of different individuals vary greatly, so it is difficult to find a unified analytical framework. Finally, corporate finance interacts with consumer finance through financial markets; consumers eventually allocate wealth in the financial market through investment, financing, savings, and so on.

According to the life cycle model proposed by Ando and Modigliani (1963), consumers will achieve wealth balance throughout their whole life cycle, and investment and financing are important tools for cross-period equilibrium. Combined with corporate finance, there are a lot of studies that discuss the investment behavior of consumers; but it is rare to find any on financing behavior. Financing is an important aspect of the interaction between consumers and financial markets, where consumers play the role of demanders. Although financing may pose risks to financial institutions, it is more likely for consumers to have opportunities to change their living conditions. However, most consumers in China have a more conservative attitude towards financing, preferring to save money rather than borrow. This has seriously restricted China's economic development.

Therefore, in this book, we will introduce the financing situation, financing constraints, financing preferences, representative financing tools and credit reference system in China, and analyze some unique financing phenomena of Chinese households on the basis of survey data and financial institutions transaction data, in order to let readers have a better understanding of consumer finance in China. The book has six chapters, which are summarized below.

Chapter 1 is a brief introduction. In this chapter, we will discuss the general consumer's financing behavior, through the data from the China Household Finance Survey conducted by the China Financial Research Center, Tsinghua University, in 2011. We will describe the current status of loans and financing; through group analysis, we can recognize some basic characteristics of the microfinance market, which is important to the prediction of individual behaviors and the expanding financing markets.

We will discuss financing constraints in Chapter 2. This is a really big problem in China; many consumers who need money can't borrow enough without adequate collateral. But financing constraints cannot be determined through direct observation, so how to measure financing constraints is a very crucial consideration. In the study, we used the data from the China Household Finance Survey conducted by the China Financial Research Center, Tsinghua University, in 2011, selected the key factors of financing constraints, and built a Probit regression model to calculate the financing constrains which a family may face. And through the analysis, we separated financing constraints into two categories: one is "reasonable financing constraints", the other is "unreasonable". So we paid more attention to the unreasonable financing constraints, and through empirical analyses, we found that bad credit records and health conditions may be the two main factors which affected consumer financing constraints—consumers who have problems with financing constraints will directly change their consuming and investing behavior. Finally, we will propose that an effective way to solve the consumer financing constraints in China is to give more trust to the new consumer and be more tolerant with history credit records.

Chapter 3 will discuss financing decisions. According to the hypothesis of rational people, people always pursue the maximization of their preferences; and this corresponds to the maximization of expected utility in risk decision-making. Based on this, we will discuss the basic framework of financing decision-making analysis, and following the framework we use the analytic hierarchy process to quantitatively analyze the decision-making process of household financing. Through the analysis, we can see two key factors—costs and convenience—are the main factors which affect financing decisions. Then we will discuss the financing decision difference among savvy groups. Finally, we will describe the financing behavior of Chinese consumers through the actual survey data, and we will conduct an empirical analysis from four factors to analyze the financing behavior. It can be seen that wealth, household structure and financial plan play important roles in long-term loan owning. And health status plays the key role in non-bank financing.

In recent years, credit cards as payment instruments and financing channels have been used to a large extent in China. Using the credit card data of the Industrial and Commercial Bank of China, we will give more details about who used credit cards as payment instruments and who used credit cards as financing channels in Chapter 4. Through the data analysis, we can find the possibility to increase credit card profit by extending some advertisement to potential consumers, giving coupons to regular customers and making applications less restrictive. So, this work can give some useful advice for credit card companies.

Chapter 5 will focus on internet financing. The advance of internet technology provides a convenient market platform for matching lending and borrowing parties, but many consumers still hesitate to use online borrowing. To better understand consumer behavior in online borrowing, we will use nationally representative survey data in China to explore factors affecting consumer use of one type of online borrowing. Through empirical analyses, we found that financing knowledge and risk attitude are two key factors associated with P2P (person-to-person) borrowing. And then we go further with the transaction data of Credit Ease; we will explore the question of who will lend on the internet, and who will get money successfully through the internet. Through the empirical analysis, we found that different regions have different markets in internet financing; but generally speaking, there is still a large space in second-tier cities for internet financing.

The last chapter will focus on the credit reference system in China. It is known that more and more data is collected in the credit reference system in China, which is the central credit information system of the country. How to classify users and conduct analysis of users' personas is becoming a big problem in practice. In our study, we will give a brief introduction of the development of the credit reference system and the credit industry. Firstly, we will provide a general discussion of the demands of data use. Through empirical analysis, we found that the gap between the data supply and the inquiry demand is now getting larger and larger. And then with sample data coming from that system, we did a cluster of consumers. Through the cluster, we gave a preliminary user portrait in the personal credit reference system. Furthermore, we will provide a discussion on the privacy protection and information sharing. Finally, we will suggest that it is better for the authorities to add more institutions in the system and share the data with non-commercial institutions for conducting more research.

This book is for a wide audience. It is suitable for companies that want to expand their business in China, it may be interesting for researchers who hope to compare some results in different countries, and it will also give some people from different countries an opportunity to know more about China's household and financial market. In short, it is for anyone who has interests in consumer financing.

The aim of the book is to give a general picture of consumer financing in China, through the survey data and transaction data; we try to find the basic factors to reach the broader conclusions. We hope that you find the book is full of useful information, and we hope the book at least enables you to enjoy consumer finance. If there is some mistyping, please feel free to contact: hanluivy@126.com.

1. Brief Introduction

In the past five years, China's average annual growth rate of wealth has been 20%. As a result, the fortune of middle-class households (with total household assets between 200,000 RMB and 500,000 RMB, as defined by the China National Bureau of Statistics) has increased at a rate of roughly 15% per year since 2012.

Meanwhile, various types of consumer financial markets have been expanding; besides housing loans, car loans and credit cards, more financing products came into our lives, such as consuming loans, education loans and venture loans. And a new financing method called a person-to-person (P2P) loan is now taking more and more attention.

The consumer financing market, as an important way for consumers to allocate wealth throughout their lifetimes, is increasingly used by middle-class families. Therefore, by investigating the household financing status, we can significantly grow our understanding of the microfinance market, which is the most important factor in predicting individual behaviors, and it will give us an outline of the development in new financing markets.

We are committed to using the basic methods of data analysis to study the basic situation of household financing in China. With the survey data of the China Household Finance Survey conducted by the China Financial Research Center, Tsinghua University, in 2011, we aim to plot a global picture of China's credit market. For more information about the survey, please see: https://www.weiyangx.com/jtxfjrsj.

1.1 Current status of loans

This study uses the data from the China Household Finance Survey conducted by the China Financial Research Center, Tsinghua University, in 2011. The sample covers 25 cities and the total number of samples is 5,911. Combined with the variables selected by the research, the number of valid samples is 4,711 after removing missing values, outliers and invalid values.

This questionnaire divides the cities in the whole country (excluding Hong Kong, Macao and Taiwan) into seven geographic regions: Northeast China, North China, East China, South China, Central China, Southwest China and Northwest China. The number of sample households in each region is selected according to the proportion of regional family households in the total number of family households. In each of the selected cities, the researchers use the random sampling method to choose samples. So, it can be seen that the samples can effectively represent the national situation.

Household loans can be divided into two sorts by time limit: long-term loans, with a term of more than one year, and short-term loans, with a term of less than one year. The long-term loans of households mainly refer to debt. Usually, households raise long-term loans from relatives, friends and banks, with the purposes of purchasing a house, a car, further education, etc. The short-term loans of households mainly refer to consuming loans. Usually, households raise short-term loans from banks and credit card companies with the purposes of consuming.

According to the life cycle model proposed by Ando and Modigliani (1963), demographic variables generally refer to the basic characteristics of a person, including gender, age, marital status, whether they have children, life cycle stage, etc. The life cycle hypothesis is the basic theoretical framework of household consumption and credit behavior, which assumes that rational consumers aim to maximize the utility of their whole life. The theory holds that in spite of the constant changes of personal income, families tend to apply financial instruments to achieve a stable consumption flow in the life cycle and the income consumption ratio is unchanged. Therefore, the life cycle hypothesis is often employed as an important basis for classifying household groups. Following this theory, we explored the status of household loans by groups. We explored the value proportion of loans with age groups, martial groups, family status groups, education groups, occupation groups, income groups and geography groups.

1.1.1 Age groups

From the age groups, households in the over-50 age group are less likely to have loans; the average amount is about 30,594 RMB in total. Households with more loans are mainly concentrated in the 25–34 age group and 35–40 age group, and the average monetary value of loans is quite high; accordingly, the means of these two groups are 259,066 RMB and 330,547 RMB in total.

Long-term loans are the main loan part in every age group, accounting for more than 70% of the total loans. Among long-term loans, the value proportion of housing loans is obviously different; it is 4.13% in the below-25 age group, 31.11% in the 25–34 age group, 46.98% in the 35–40 age group, 9.12% in the 41–50 age group, 3.81% in the 51–60 age group and 1.04% in the over-60 age group. Also, the value proportion of car loans is somewhat different: 2.38% in the below-25 age group, 48.22% in the 25–34 age group, 20.56% in the 35–40 age group, 3.18% in the 41–50 age group, and none in other age groups. Other average monetary value of long-term loans is relatively small, so we do not report their proportions in detail. From these, we can see that to buy a house is the major loan purpose for a family. And the occupancy rate of housing loans is quite high; nearly half of households in this survey have housing loans. No matter the age group of the household, the proportion of households who borrow to buy a car is not very high; the figure for the highest group (35–40 age group) is only 6.17%. The value proportions among age groups can be found in Figure 1-1.

The number of people who have short-term loans is generally low. Relatively speaking, households in the below-25 age group and the 25–35 age group are the majority people owning short-term loans. What is more, households with occupants under age 25 have more amount proportion of consuming loans than other groups. And households with occupants over the age of 50 don’t hold any short-term loans, which is different from other groups. Besides this, the proportion of households who have short-term loans is about 8%, not above 10%.

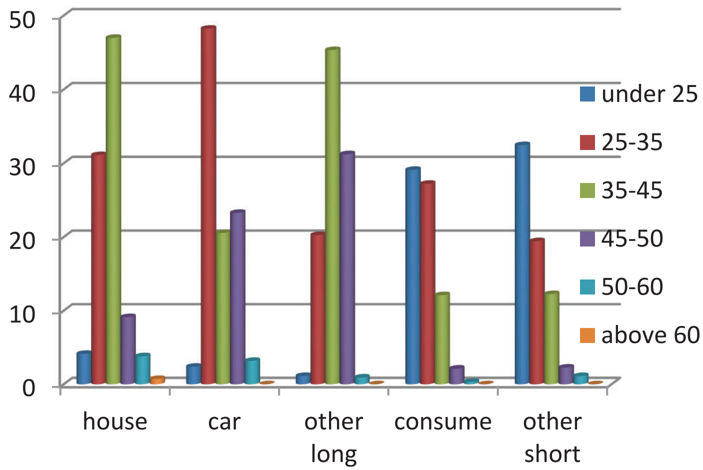

Fig 1-1 Value Potation of Loans between Age Groups

1.1.2 Marital groups

From the marital groups, the married families are more likely to hold loans, with a mean of 150,789 RMB, compared with the unmarried families, who have a loan mean of about 43,421 RMB.

Long-term loans form the main part in every group; they account for more than 60% of the total loans. Among long-term loans, the value proportion of housing loans is obviously different: it is about 76.84% in the married group and 23.16% in the unmarried group. Also, the value proportion of car loans is relatively different; it is 81.36% in the married group and 18.63% in the unmarried group. Other monetary amounts of long-term loans are also quite different: about 80.79% in the married group and 19.22% in the unmarried group. From these, we can see that married families are more likely to have long-term loans. And the proportion of long-term loans in married families is over 15%, which means that in our survey over 15% of married families have long-term loans, compared with the unmarried families, who have a rate of only 8%. The value proportion among marital groups can be found in Figure 1-2.

Short-term loans are not the main part in every group; they only account for nearly 15% in total. Among short-term loans, the value proportion of consuming loans is obviously different, though the main part of consuming is still within married families; the unmarried group borrows 33.37% in total value, which gives consuming loans the smallest gap between the two groups. Also, the value proportions of other short-term loans are relatively different; the value proportion is 88.34% in the married group and 11.65% in the unmarried group, From these, we can see that unmarried families are more likely to have consuming loans. The proportion of short-term loans is about 8.74%, which means that in our survey 8.74% of unmarried families have consuming loans, compared with the married families, whose rate is only 7.88%. The value proportions among marital groups can be found in Figure 1-2.

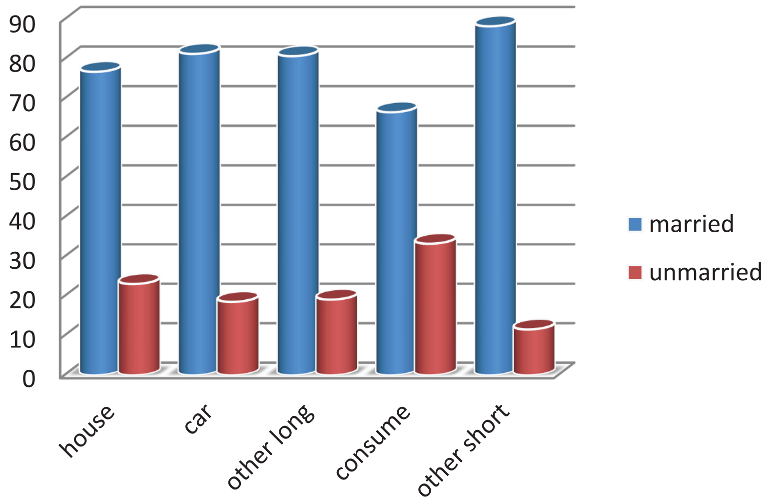

Fig 1-2 Value Potation of Loans between Marital Groups

1.1.3 Family status groups

From the family status groups, we take the number of children and the number of elderly people into consideration. In total, it can be seen that the households without children or with one child are more likely to hold long-term loans; the mean of those are 80,713 RMB and 102,441 RMB, respectively. Generally speaking, for the number of elderly people, it can be seen that households with more elderly people will be more likely to hold short-term loans; the means of the five groups (none elderly, one elderly, two elderly, three elderly, and four elderly or more) are 5,120 RMB, 8,335 RMB, 8,917 RMB and 9,633 RMB.

Loan-term loans are the main part in every group, accounting for more than 50% of the total loans. Among long-term loans, the value proportion of housing loans is obviously different; it is about 42.68% in the no-child group and 44.12% in the one-child group, and these two groups are the main holders of loans. Similarly, the value proportions of car loans are relatively different; the value proportion is 52.83% in the one-child group and 32.93% in the no-child group. Other monetary amounts of long-term loans are also the same; that is, about 50.33% in the one-child group and 34.52% in the no-child group. From these, we can see that no-child and one-child families are more likely to have long-term loans. Besides these statistics, the proportions of these two groups are not so large, just about 15%; that means that in our survey, 15% of the households with one child and no child have long-term loans, compared with other groups, whose rates are about 20%. Also, for the elderly groups, the value proportions of housing loans are different among the groups; the groups with two elderly people and four or more elderly people have larger proportions, which are 30.86% and 34.26%, respectively, just the same as car loans and other long-term loans. The value proportions among family status groups can be found in Figure 1-3 and Figure 1-4.

In terms of short-term loans, families without children account for more than 50% of the total consuming loans. Families with one child have the largest proportion of other short-term loans, accounting for 69.2%. The difference is that consuming loans are relatively high in households with two or four or more elderly people; meanwhile, other short-term loans account for 33.93% of the households with four or more elderly people. From these numbers, we can infer that no-child and one-child families are more likely to have consuming loans. At the same time, the consuming loan rates of these two groups are very high, up to 50%, which means that in our survey over 50% of the households with one child and no child have consuming loans. Also, families that support two or four or more elderly people will have more consuming loans, and the rates of consuming loans in these two groups are relatively high, up to 20%. The value proportions among family status groups can be found in Figure 1-3 and Figure 1-4.

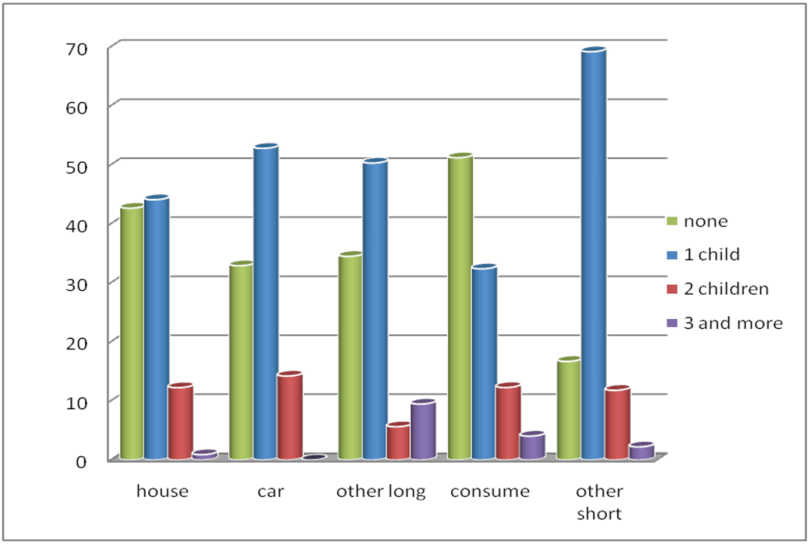

Fig 1-3 Value Potation of Loans between Family Status Groups with Children

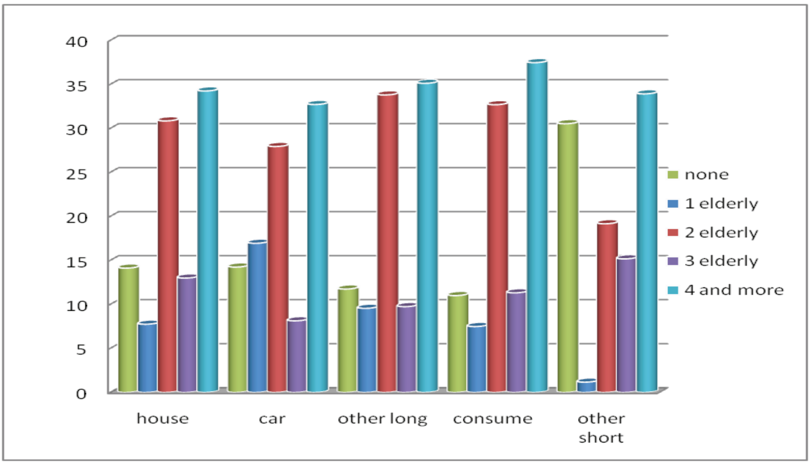

Fig 1-4 Value Potation of Loans between Family Status Groups with Elderly People

1.1.4 Education groups

From the education groups, the undergraduate group is more likely to hold loans, with a mean of 201,634 RMB, compared with the high-school-and-below group, which has a loan mean of about 83,421 RMB, and the mean in the graduate group, which is about 145,388 RMB.

In terms of long-term loans, the total distribution of housing loans is more than 60%. There are obvious differences among different education groups. The undergraduate group accounts for 61.31% of the total housing loans, while the other two groups account for relatively low proportions, which are 31.8% with the high-school-and-below group and 6.88% with the graduate group. Similarly, the graduate group also holds more car loans and other long-term loans. At the same time, we also note that in the survey, more than 30% of undergraduates will have long-term loans, while the rates of the other two groups who have long-term loans are 13% and 48%. The value proportion among education groups can be found in Figure 1-5.

In terms of short-term loans, the total distribution of consuming loans is about 53%. There are obvious differences among different education groups. The undergraduate group accounts for 64.6% of the total consuming loans, while the graduate group accounts for a relatively low proportion, which is only 3.93%. However, the high-school-and-below group holds more other short-term loans—51.68% in total. What is more, other short-term loans are seldomly held by the graduate group; the percentage of these loans is only 0.21%. At the same time, we also note that in the survey, nearly 25% of the graduate group members will have short-term loans, while the rates of the other two groups who have short-term loans are about 4% and 6%. The value proportions among education groups can be found in Figure 1-5.

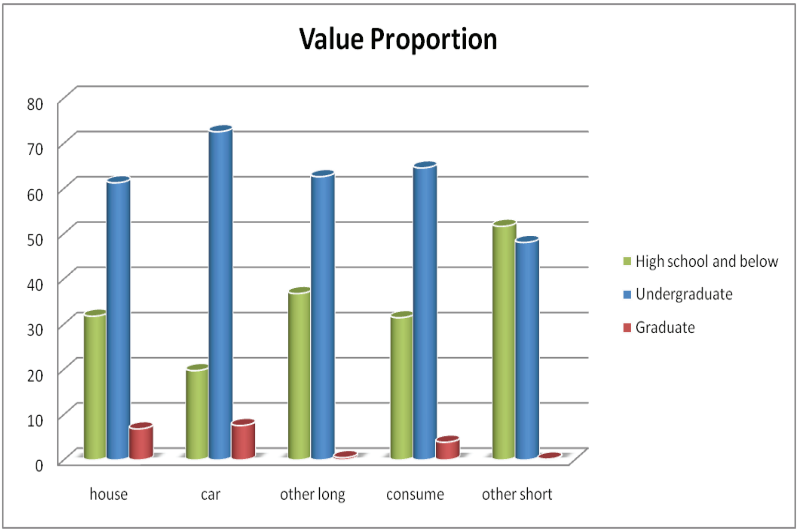

Fig 1-5 Value Potation of Loans between Education Groups

1.1.5 Occupation groups

From the occupation groups, the enterprise group is more likely to hold loans, with a mean of 254,821 RMB, compared with the unemployed group, which has a loan mean of about 29,347 RMB.

In terms of long-term loans, the total distribution of housing loans is more than 40%, and then the major part is car loans. There are obvious differences among different occupation groups. The enterprise group accounts for 48.79% of the total housing loans, and the self-employed group accounts for 24.54% of the total housing loans, while the other groups account for a relatively low proportion. The retirement group has the smallest number of people with housing loans (about 1.97%). The government and unemployed groups also have low levels of housing loans: 3.93% of the government group and 2.01% of the unemployed group. Similarly, the enterprise group and self-employed group also hold more car loans and other long-term loans, while the unemployed group and retirement group have the least. At the same time, we also note that in the survey, more than 25% of households in the enterprise group and self-employed group will have long-term loans, while the rates of other groups who have long-term loans are 14.13% in the unemployed group and 6.06% in the retirement group.

The value proportion among occupation groups can be found in Figure 1-6.

In terms of short-term loans, the total distribution of consuming loans is about 62%. There are obvious differences among different occupation groups. The enterprise group accounts for 48.38% of the total consuming loans, and the self-employed group accounts for 25.13% of the total consuming loans, while the unemployed group accounts for the lowest proportion, which is only 3.71%. However, for the other short-term loans, the self-employed group takes the biggest part, which is about 47.62%, while the unemployed group takes the smallest proportion—about 1.19%. At the same time, we also note that in the survey, about 10% of the self-employed group will have short-term loans; it is much more than other groups. The value proportion among occupation groups can be found in Figure 1-6.

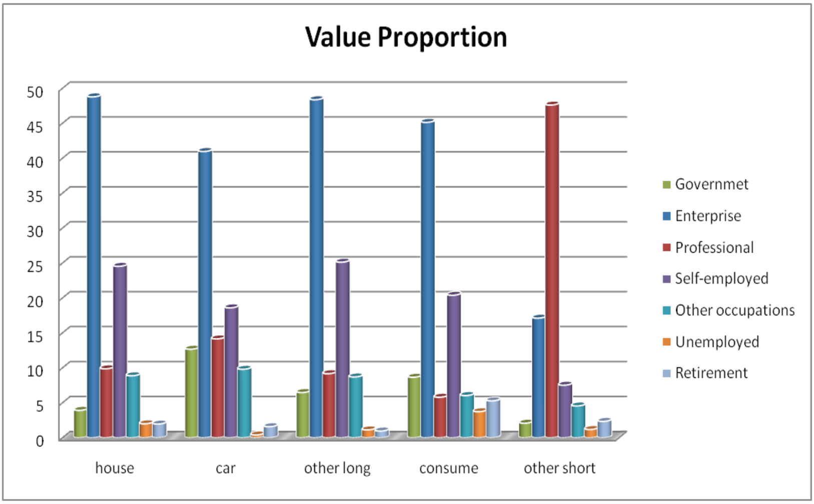

Fig 1-6 Value Potation of Loans between Occupation Groups

1.1.6 Income groups

According to the division method of the National Bureau of Statistics, income groups of family annual income can be divided into seven grades which are (in RMB): below 10,000; 10,000–20,000; 20,000–50,000; 50,000–100,000; 100,000–200,000; 200,000–500,000;

and above 500,000.We use the grade numbers 1 to 7 for short. It can be seen that with the increase in grade, more and more households will have loans. The average loan amounts of these seven groups are, respectively (in RMB): 2,312; 7,552;8,404; 25,360; 45,703; 80,027; and 208,305.

The divisions of long-term loans in this category are quite different. Long-term loans, which form the main part of loans for groups 1, 2, 3 and 5, vary obviously, and the proportion of total long-term loans ranged from 20% to 80% with significant differences. Among long-term loans, the value proportions of housing loans are obviously different; groups 4 and 5 have the largest number of mortgages, with value ratios of about 28.49% and 27.36%, respectively. Similarly, car loans account for more than 20% in groups 4, 5 and 7; respectively, the ratios are 26.21%, 24.97% and 20.56%. Group 4 has the highest value ratios for other long-term loans, exceeding 38.41%. And more, we also find that households with annual incomes of 50,000–500,000 RMB are more likely to have long-term loans, and the proportion of these is over 20%. That means, in our survey, over 20% of the families with 50,000–500,000 RMB annual income have long-term loans; by comparison, in the other groups this rate is no more than 10%. The value proportion among income groups can be found in Figure 1-7.

The distribution of short-term loans varies greatly among groups. Generally speaking, groups 4, 5 and 7 have relatively more short-term loans. Among short-term loans, the value proportions of consuming loans are obviously different; groups 4 and 5 have the largest numbers of consuming loans, with value ratios of about 27.55%, 23.14% and 22.18%, respectively. Group 7 has the most other short-term loans in value, exceeding 34.54%. What is more, we also find that households with annual incomes over 500,000 are more likely to have short-term loans, and the proportion of these is over 10%, which is much lower than the proportion of long-term loans. This means that in our survey over 10% of household whose annual income over 500,000 have short-term loans; by comparison, n the other groups this rate is no more than 5%. The value proportion among income groups can be found in Figure 1-7.

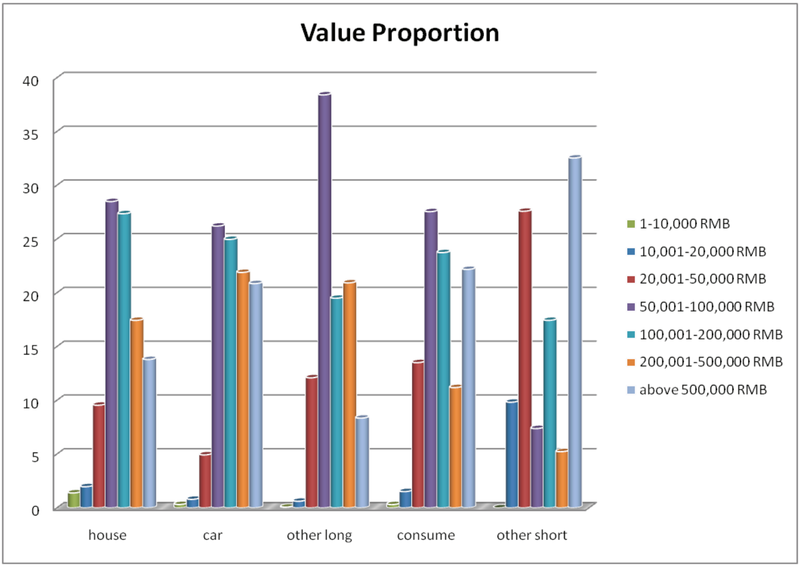

Fig 1-7 Value Potation of Loans between Income Groups

1.1.7 Geography groups

Overall, households in the first-tier city have more loans, and the mean of these is 118,467 RMB; by comparison, in the second-tier city the loan mean is about 35,824 RMB, and it is 48,726 RMB in the third-tier city.

Long-term loans are the main part in every group, accounting for more than 60% of the total loans. Among long-term loans, the value proportion of housing loans appears to be characteristic of U-type distribution: about 55.58% in the first-tier city, 10.71% in the second-tier city and 33.71% in the third-tier city. Also, the value proportion of car loans is U-shaped: 49.58% in the first-tier city, 9.36% in the second-tier city and 33.71% in the third-tier city. Similarly, the monetary amount of other long-term loans is also U-shaped: 41.75% in the first-tier city, 5.19% in the second-tier city and 53.06% in the third-tier city. What is more, we can see that households in the third-tier city are more likely to have loans, and the proportion of these is over 60%.That means that, in our survey, over 60% of families in the third-tier city have long-term loans, compared with the first-tier city and second-tier city, where the rates are nearly 20% and above 10%, respectively. The value proportion among geography groups can be found in Figure 1-8.

Short-term loans have different distributions. Among short-term loans, the value proportion of consuming loans appears to be characteristic of U-type distribution: about 56.53% in the first-tier city, 5.05% in the second-tier city and 38.42% in the third-tier city. In other short-term loans, the distributions are more concentrated. The largest part is in the third-tier city, which accounts for 76.07%. However, we can also see that there is no significant difference among the three groups in the holding rate, which is less than 5%. The value proportion among geography groups can be found in Figure 1-8.

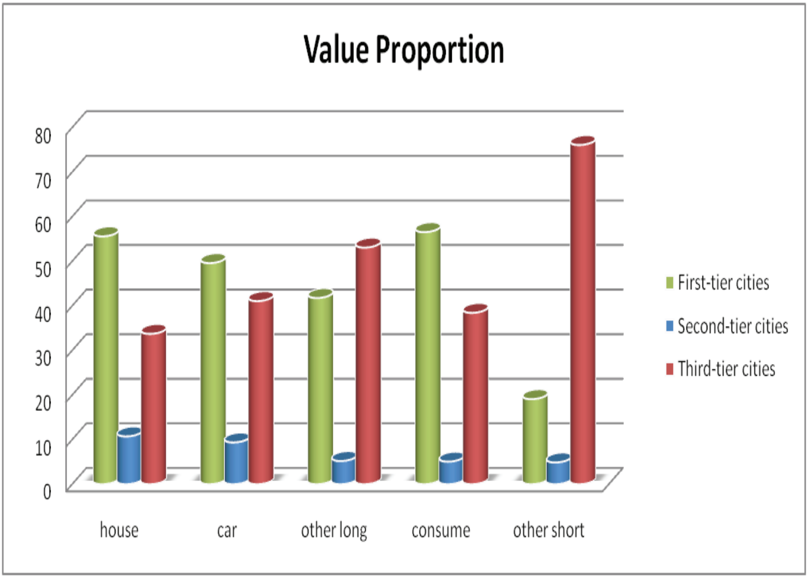

Fig 1-8 Value Potation of Loans between Geography Groups

1.2 Current status of financing

There are four main channels for personal financing in China, which are: banks, non-bank institutions, relatives or friends, and private lending. Here we summarize three aspects of the current situation: financing preference, financing knowledge and financing affordability.

1.2.1 Financing preference

Overall, households in China prefer to borrow from relatives or friends. We summarize the financing preference for household wealth in Table 1-1.

Table 1-1 Preferred financing channels for household wealth

(Unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
Relatives or friends	76.22%	87.66%	74.40%	78.84%	71.03%	72.30%	67.79%
Banks	22.10%	10.35%	24.35%	18.15%	24.02%	25.90%	29.97%
Non-bank institutions	0.55%	1.45%	0.42%	1.45%	3.80%	0.62%	1.11%
Others	1.13%	0.55%	0.83%	1.56%	1.16%	1.17%	1.13%

Figure 1-9 also reveals the preferred borrowers in China, but in Figure 1-9 we put our focus on the number of borrowers. As can be seen, most households in China prefer to finance from relatives and friends. And the less wealth a family has, the more preference they have; and as the family wealth rises, the proportion of financing from banks gradually increases.

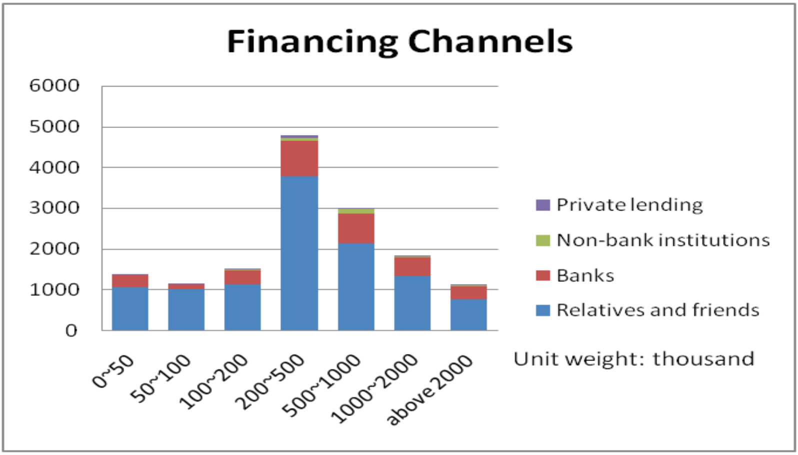

Fig 1-9 Household financing channel preferences

Further, we have explored a relatively special financing channel in China at this stage, which is the financing from relatives and friends. Usually, this financing channel does not require collateral or even interest, and it is based entirely on the trust relationship between each person. Figure 1-10 depicts the distribution of interest on borrowing from relatives and friends, and Figure 1-11 depicts how families get financing from relatives and friends.

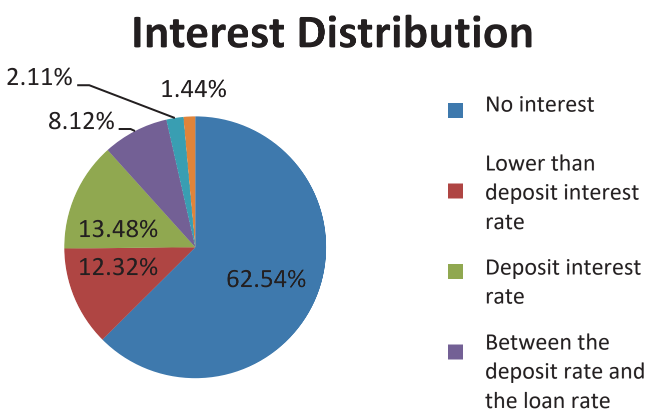

Fig 1-10 Interest distribution of financing from relatives or friends

From Figure 1-10, we can see that more than 60% of financing from relatives and friends is without interest; only 3.55% of these financing interest amounts will exceed the direct loan rates from financial institutions.

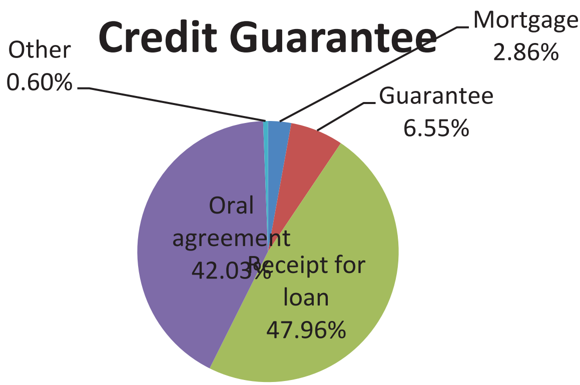

Fig 1-11 Credit guarantees for financing from relatives or friends

Figure 1.11 depicts the ways in which Chinese households borrow money from relatives and friends. Most residents (47.96%) borrow money from relatives and friends in the form of an IOU. This is followed by oral agreement, with a ratio of 42.03%. A further 6.55% of the residents will ask a middleman to guarantee, and 2.86% will use their mortgage to borrow money from relatives and friends.

1.2.2 Financing knowledge

We can divide the financing products into six categories according to the purpose of the financing: house loan, car loan, consuming loan, venture loan, education loan, decoration loan and so on. We further summarize the knowledge degrees about these six categories according to family wealth status.

Tables1-2 to 1-7 reveal the knowledge degrees of households in China regarding house loans, car loans, decoration loans, education loans, venture loans and consuming loans. It can be seen that with the increase of the total household wealth, the level of knowledge of these loans has also increased.

Table 1-2 Knowledge degree of house loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
1	53.37%	52.59%	51.07%	45.21%	21.97%	24.55%	16.34%
2	23.01%	19.47%	18.26%	29.23%	27.64%	21.96%	15.61%
3	19.38%	24.68%	26.92%	20.62%	37.72%	35.82%	43.33%
4	3.73%	3.15%	3.21%	4.24%	10.06%	14.24%	16.36%
5	0.52%	0.11%	0.54%	0.71%	2.61%	3.43%	8.35%

Note: the knowledge degree is self-assessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

Table 1-3 Knowledge degree of car loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
1	64.40%	55.99%	53.22%	47.98%	29.67%	29.64%	22.68%
2	20.70%	26.00%	27.30%	33.94%	31.84%	24.67%	25.47%
3	12.05%	17.08%	17.43%	14.56%	29.76%	35.83%	34.39%
4	2.67%	0.87%	1.94%	2.99%	6.89%	8.22%	14.39%
5	0.17%	0.06%	0.12%	0.53%	1.84%	1.64%	3.08%

Note: the knowledge degree is self-accessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

Table 1-4 Knowledge degree of decoration loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
1	68.57%	63.65%	60.77%	60.07%	40.88%	41.74%	33.29%
2	22.24%	26.68%	27.56%	30.03%	41.90%	33.39%	36.24%
3	6.27%	9.27%	9.45%	7.34%	12.21%	20.80%	21.83%
4	2.74%	0.29%	2.10%	2.42%	3.63%	2.92%	6.11%
5	0.17%	0.11%	0.12%	0.14%	1.38%	1.14%	2.54%

Note: the knowledge degree is self-assessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

Table 1-5 Knowledge degree of education loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
1	64.77%	63.38%	56.79%	52.76%	37.71%	40.17%	37.22%
2	22.03%	22.77%	19.87%	26.54%	34.83%	30.39%	33.54%
3	9.12%	12.21%	20.02%	15.58%	21.00%	22.28%	22.78%
4	3.73%	1.37%	3.28%	4.40%	4.72%	6.48%	4.50%
5	0.36%	0.27%	0.04%	0.72%	1.73%	0.68%	1.97%

Note: the knowledge degree is self-accessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

Table 1-6 Knowledge degree of venture loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**
1	68.13%	66.57%	63.78%	55.66%	39.69%	38.19%
2	19.87%	25.07%	22.41%	30.36%	36.42%	30.18%
3	7.69%	8.01%	11.08%	10.44%	17.96%	25.14%
4	2.80%	0.35%	2.60%	3.01%	3.66%	5.14%
5	1.51%	0.00%	0.12%	0.53%	2.27%	1.35%

Note: the knowledge degree is self-assessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

Table 1-7 Knowledge degree of consuming loans

(unit: thousand yuan)	**0–50**	**50–100**	**100–200**	**200–500**	**500–1,000**	**1,000–2,000**	**2,000+**
1	72.77%	69.48%	66.20%	60.14%	44.14%	44.98%	38.79%
2	17.10%	22.62%	20.99%	26.83%	35.55%	33.74%	31.29%
3	7.14%	6.93%	10.59%	9.31%	16.63%	17.82%	21.30%
4	2.12%	0.97%	2.17%	2.03%	2.45%	2.46%	5.31%
5	0.86%	0.00%	0.04%	1.69%	1.23%	1.00%	3.31%

Note: the knowledge degree is self-accessed; 1 refers to having no knowledge, and 5 refers to knowing very well.

From the results of the data statistics, we can see that most households do not know much about the loan products. They are more familiar with house loans and car loans to some degree; however, most households think that they do not know enough about venture loans and education loans. Relatively speaking, most households have knowledge of house loans and car loans, and only some households with high financial status have a better understanding of venture loans. This may be the key point which affects their behavior in financing.

1.2.3 Financing affordability

We measure household financing affordability by the ratio of total household loan to annual household income. We assume that the loan could be taken on as a multiplier of the family's annual income, which can be divided into 10 levels, such as less than 1 time, 1 time, 2 times... 10 times, etc. And we explored the financing affordability distribution from total family wealth as shown in Figure 1-12 and risk aversion as shown in Figure 1-13.

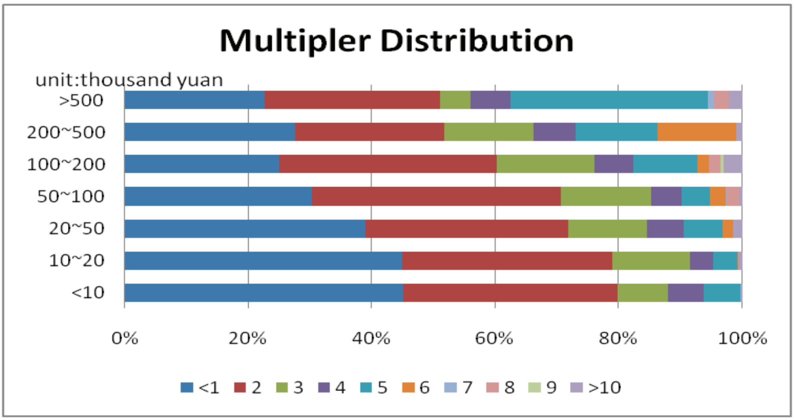

Fig 1-12 Multiplier distribution by household wealth

From Figure 1-12, we can see that with the increase of household wealth, the loan multiplier acceptable to households has a significant growth trend. However, most households consider the most acceptable loan multiplier to be less than three times their annual income.

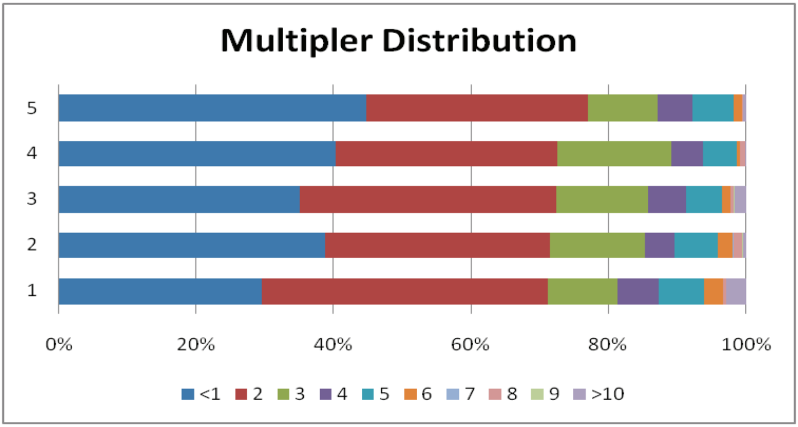

Fig 1-13 Multiplier distribution by household risk attitude

We divide family risk attitudes into five groups, from low to high risk. Preference increases in turn: 1 represents risk aversion and 5 represents risk preference. From Figure 1-13, we can see that with the increase of household risk preference, the loan multiplier acceptable to households has a significant growth trend. However, most households consider the most acceptable loan multiplier to be under twice their annual income, no matter which risk attitude group they are.

Most families (62.54%) will not pay interest. The second group, 13.48%, pays according to the deposit interest rate. Again, the third group pays higher than the deposit interest rate, the ratio is 12.32%. Only 2.11% of residents will pay according to the loan interest rate, and 1.44% of the residents will pay higher than the loan interest rate.

2. Financing Constraint

Financing constraint is a big problem which many households may face. Duygan-Bump, Levkov, and Montoriol-Garriga (2015) showed that financing constraints of small firms owned by household faced during the 2007–2009 recession in the United States. Ruiz-Tagle and Vella (2016) argued that financing constraint had a big effect on consumer credit behavior.

In addition, financing constraint cannot be discerned through direct observation, so the question of how to measure financing constraints is very crucial. Due to different research questions or different data used, scholars at home and abroad measure financing constraints in different ways. The main idea is that if consumers have financing needs but are not fully satisfied, they are subject to credit constraints. Duca and Rosenthal (1993) put consumers in the category of "facing credit constraints" using the measure from financing institutions; they measured the loans which are applied but not approved as supply constraints and which household are not willing to use their credits as demand constraints. This is known as a supply-side financing constraint. Grant (2007) studied the credit constraints faced by American households; it is considered that families have financing needs and are subject to credit constraints when demand is greater than supply. This is known as a "demand-side financing constraint". So, we put our attention into how to measure the credit constraint.

Credit constraints will affect consumer behavior. Getachew (2016) examines how credit constraints affect the dynamics of wealth and thereby the dynamics of capital and output growth. Hai and Heckman (2017) found evidence of substantial life cycle credit constraints that affect human capital accumulation and inequality. Li, Lin, and Gan (2016) explored the impact of credit constraints on rural households' consumption expenditure in South China, and they suggest that "relaxing the credit constraints helps to improve the rural households' consumption expenditures in developing countries". But there is little research which discussed the urban household credit constraints, and they affect one's behavior, especially the consuming behavior and investment behavior.

In this chapter, we first describe the credit constraints of Chinese households. Through empirical research, we found the main factors that may lead to credit constraints, which are household size, credit card record and income. Furthermore, we discuss the consumption and investment behavior of credit constraints groups. It can be found that under credit constraints, household consumption is significantly inhibited, and household investment is more inclined to real estate investment.

2.1 Measure of credit constraint

2.1.1 Relevant research

In the research of credit constraint, a popular view claimed an explanation for rejecting the hypothesis which is put forward by Hall (1978). The view claimed that at least some consumers face binding credit constraints. Because of this, some cannot borrow as much as they want at "market liquidation" rates. Some typical research can be found in the work of Jaffee and Russell (1976), Stiglitz and Weiss (1981), Hayashi (1987), Kehoe and Levine (1993) and Kocherlakota (1996). Hayashi defined two types of credit constraint, one is that consumers can not borrow enough amount; the other is the borrowing interest is much higher than the market liquidation rates. The first type is also called as credit rationing. In the recent research, researchers put more attention on credit rationing. Generally speaking, if the benefits of default outweigh the penalties, the borrower will default. By solving the decentralized market economy without family default in equilibrium, these models endogenously stimulate credit rationing, that is, borrowing is limited to a maximum level depending on the parameters of the models. However, these models treat the interest rate as the same.

In our research, we focus on the measure of consumer financing constraint. Using 2011 China Household Finance Survey micro-data, we explore the current consumer financing constraint situation in China, and then we discuss the roles which may impact credit constraint. This chapter is organized into four sections. The first section introduces the problem which we researched and summarizes the related literature; the second section describes the data and model; the third section is the empirical results; and the last section is the conclusion.

Understanding how households' consumption responds to a credit crunch has been a central goal of macroeconomics. Alonso (2018) believes that most of the recent research has explored this question using a "hard constraint" modeling device, where households can borrow at the risk-free rate only up to an exogenous amount. An alternative, and more realistic, way to model financial frictions is to allow households to borrow as much as they want but at an interest rate that depends on the level of debt. He refers to the latter as the "soft constraint" model. He calibrated two economies differing only in the type of financial constraint that households face, and I show that a credit crunch in the hard-constraint economy (i.e. decrease in the exogenous borrowing limit) produces a drop in consumption significantly more severe than an equivalent crunch in the soft-constraint version.

In recent years, more and more studies on the current situation and expansion of credit constraints have been explored. Feng, Mu, Hu, and Kumar (2014) explored the influencing factors of family credit constraints in China using a Probit model, and empirical results show that the probability of credit constraints for family is influenced by factors like family net wealth, income, housing, age, education level, household registration and loans from other channels. Banerji, Raj, and Sen (2016) found that income, project size and household population have a significant positive impact on the likelihood of farmers getting credit when studying the influencing factors of household credit constraints. Kim, Wilmarth, and Choi (2016) emphasized that current income, wealth and age are the most important factors that determine whether consumers are facing credit constraints. Tavares-Gärtner, Pereira, and Brandão (2018) explored Contingent Payment Mechanisms (also known as Contingent Earn-Outs) as a capital-raising strategy surrounding entrepreneurial financing decisions; through their research, they argued that credit constraint has an important impact on contingent payment.

Meanwhile, credit constraint has an important effect on human daily life and financial behavior. Li, Lin, and Gan (2016) explore the impact of credit constraints on rural households' consumption expenditure in South China. Previous studies have ignored the endogeneity between the credit constraints and consumption expenditure. They use two instrumental variables to resolve this problem. Their research shows 54.9% of the respondents are credit constrained. The instrument variable model results reveal that the percentage of rural households whose consumption expenditures are credit constrained is 7.34% less than the percentage of those who are not credit constrained. Hai and Heckman (2017) investigate the determinants of inequality in human capital with an emphasis on the role of the credit constraints. We develop and estimate a model in which individuals face uninsured human capital risks and invest in education, acquire work experience, accumulate assets and smooth consumption. Similarly, Cai, Song, Ma, Dong, and Xu (2018) evaluated the impact of credit constraint on entrepreneurship by utilizing 2011 China Household Finance Survey micro-data. Hedlund (2018) investigates how equilibrium house prices respond to a tightening in credit constraints under two different but similarly calibrated models: one an infinite-horizon setting, and the other a life cycle environment.

2.2.2 Data and variables

This study uses the data from the 2011 China Household Finance Survey conducted by the China Financial Research Center, Tsinghua University. The study covers 25 cities and the total number of samples is 5,910. Combined with the variables selected by the research, the number of valid samples is 4,711 after removing missing values, outliers and invalid values.

The measurement of financing constraints in this study is designed as follows: Question (1): Did your family borrow money last year? (Excluding credit card); Question (2): Have you fully raised the required amount of money? If the participant selected “no borrowing” in question (1) and there is only partial borrowing in question (2), they are considered as being subject to credit constraints.

The final number of samples subject to financing constraints is 312, accounting for 6.62% of the total effective samples, and the data obtained is much lower than the whole samples on the status quo of credit constraints in China, mainly because most of the people responding to the first question did not borrow in the past year, which to some extent reflects the concept in Chinese families that they are not willing to borrow money to spend. Table 2-1 gives detail about the statistics regarding household financing constraints in China.

Table 2-1 Basic statistics of China's Household Financing Constraints

	Borrow	Not borrow
Number	692	4019
Proportion	14.69%	85.31%

	Did not raise money	Raised some money	Raised enough money
Number of valid samples	133	179	380
Proportion	2.82%	3.80%	8.07%

Moreover, through the survey, we also explored the consumers' feelings regarding financing constraints; in our research we described the feeling of financing constraints according to the difficulty of borrowing. In the survey, we set the question as "If you have to raise 100,000 RMB and it must be returned in one year, what do you think of this?" Figure 2-1 describes the difficulty degree in relation to household wealth, while Figure 2-2 describes the difficulty degree by income.

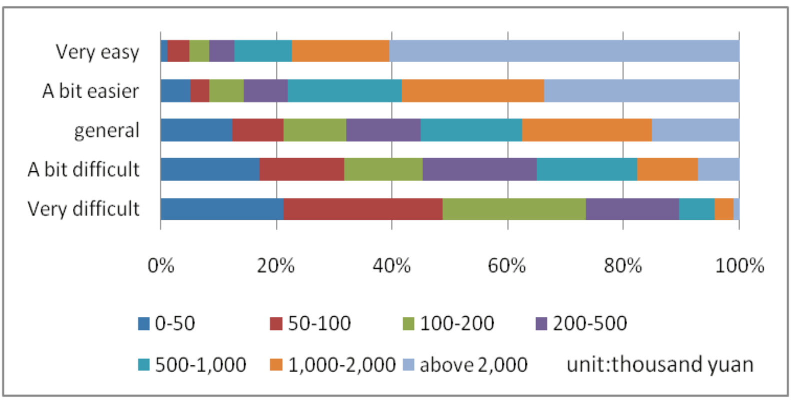

Fig 2-1 The difficulty of raising money by family wealth

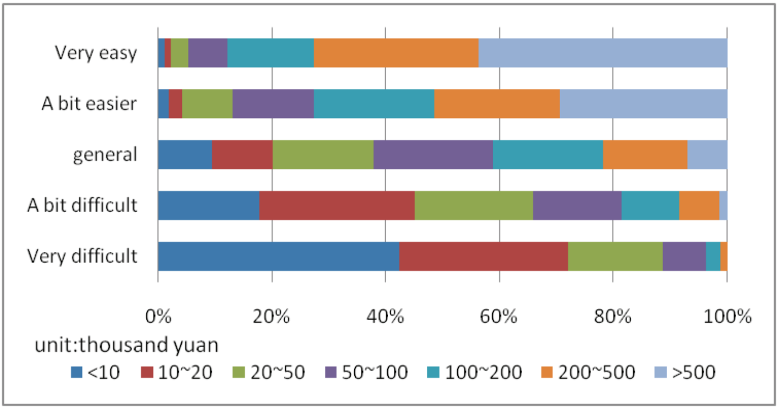

Fig 2-2 The difficulty of raising money by annual income

From Figure 2-1 and Figure 2-2, we can see that when the wealth of a household exceeds 500,000 RMB or its annual income exceeds 50,000 RMB, the participants begin to feel easy about borrowing 100,000 RMB with the maturity in one year.

Considering that the financing constraint is a discrete variable, and it can be calculated by "to raise 100,000 RMB and the maturity is one year", the household refers to "Very difficult" as the number 1, "A bit difficult" as 2, "Normal" as 3, "A bit easier" as 4, and "Very easy" as 5. And this is the dependent variable for the empirical study.

In the study of the impact factors of financing constraints, combined with previous studies, the main explanatory variables here are the marital status of the household, age, working years, education, the number of elderly people, the total number of households, the market value of the house, whether a credit card was deferred in the past year, the health of family members, the total wealth and household income in the last year. The variables are described in Table 2-2.

Table 2-2 Variable Description Table

Variable name	Variable explanation
marriage	Marital status of the head of household: unmarried = 1, married = 2
age	Age of the head of household
working years	The number of years since the head of household joined the workforce, in units of years
education	Education status of the head of household: high school and below = 1, undergraduate = 2, graduate = 3, other= 4
elderly people	The number of elderly people in the family
size	The total population of the family
house value	The market value of the house
credit card	Credit card postponement record
health	No one was hospitalized in the past three years = 1; no one was hospitalized in the past year = 2; other = 3
wealth	Total wealth
income	Total annual income

2.2.3 Empirical results

A Probit regression model is used to explore the impact factors of financing constraints. The financing constraint is expressed as follows (2.1):

$$constra = \begin{cases} 1 \\ 0 \end{cases} \tag{2.1}$$

The basic measurement model is constructed as follows (2.2), where, in the Probit model, G is the standard normal cumulative distribution function, which is $G(z) = \phi(z) = \int_{-\infty}^{z} \phi(v)dv$, where $\phi(v) = (2\pi)^{-1/2} exp((-v)^2/2)$.

$$P(constra = 1|X) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) \tag{2.2}$$

In equation (2.2), X represents the set of all explanatory variables, including marital status of the head of household (marriage), age (age), years of working (working years), education (education), the number of elderly people the family needs to support (elderly people), the total population of the family (household size), the house market value (house value), whether a credit card has been deferred in the last year (credit card), family health (health), family total wealth (wealth) and family income (income). The basic statistical analysis of the data mainly describes mean, standard deviation, minimum and maximum values of each variable. The specific results are presented in Table 2-3.

Table 2-3 Descriptive statistics table

Variable	Mean	Std. Dev.	Min	Max
marriage	1.73	0.52	1	3
age	35.59	11.09	16	83
working years	12.9	10.64	0	60
education	2.48	0.77	1	5
elderly people	1.81	1.43	0	7
size	3.43	1.54	0	30
house value	6.79	4.63	1	17
credit card	0.49	0.74	0	4
health	2	0.82	1	3
wealth	5.42	1.73	1	8
income	3.2	1.18	1	7

In this part, we use the Probit model to explore the influencing factors of Chinese household credit constraints. The specific Probit regression results are presented as Table 2-4.

Table 2-4 Probit Regression Results Table

Independent variable	Coefficient	P
marriage	0.050	0.143
age	-0.002	0.124
working years	-0.006**	0.032
education	0.029	0.216
elderly people	0.0651***	0.000
size	0.0838***	0.000
house value	0.008**	0.066
credit card	0.222***	0.000
health	0.114***	0.000
wealth	-0.018	0.213

income	-0.022***	0.001
Constant	-1.784***	0.001
Adjust R square	0.182	

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

As can be seen from Table 2-4, the number of elderly people supported by the family (elderly people), the total number of households (household size), family health (health), whether credit card payments have been deferred in the last year (credit card) and the household income in the last year (income) are significant at a 1% significance level, which means that these variables have a very significant impact on whether households are likely to face credit constraints.

Among them, elderly people, household size, credit card and health have a significant positive impact on whether the family is subject to credit constraints. This means more members, more elderly people, and more unhealthy members within a household make a family more likely to face credit constraints.

In addition, the health condition also affects the ability of borrowers to repay. Some formal financial institutions, such as banks, may not be able to determine whether a borrower can repay their loans within the stipulated time limit based on their health status. However, when borrowers get money through relatives and friends, the acquaintances will consider the borrower's ability to repay in the future and may not want to lend money to a borrower with a bad health condition, so the family's health condition will definitely affect the credit constraints on the household.

Income has negative relationship with credit constraint; with the increase of household income, household credit constraints will decrease significantly. The same relationship can be found with the working years—it can be seen that it also has a negative relationship with credit constraint. As we know, there is a positive correlation between working years and income to a degree, so we can attribute this to the increase of total income.

In the past year, the more times the family has postponed credit card payments, and the worse the family's health condition is, the more likely the family is to face credit constraints. Delay in repaying credit card loans, to a certain extent, reflects personal credit. Financial institutions think a great deal about personal credit in the issuance of personal loans, and credit will be based on personal credit history, so deferred credit card payments will affect the probability of the family successfully getting loans from formal financial institutions.

2.2.4 Conclusion

To sum up, we have a preliminary discussion on the credit constraints of urban households in China. Through the survey data, we can find that over 45% of households with borrowing behavior have financing constraints. At the same time, most families find it becomes easier to borrow money when their wealth exceeds 1 million RMB or their annual income exceeds 500,000 RMB. So, it can be inferred that credit constraint is still a big problem which Chinese households face.

In addition, we can find that bad credit records, family burden and income may be the main factors which lead to consumer financing constraints. If one has credit card payment deferring records, it may lead to credit constraints from formal institutions. If one has a relatively heavy family burden, it may lead to credit constraints from personal lending. And with the increasing of income, the credit constraints will be less. So, we can see credit constraints may be a bottleneck for the poor household. How to break credit constraints needs further discussion.

2.2 Consuming behavior under credit constraint

In the research of Bauman (2001), he claims a new form of consumption, which consists of a transformation from the demand function to plastic and volatility of dispersion. He believed that this unstable principle had become a function of modernity, which seems to be stable due to lack of rigidity, so credit financing is an effective tool of consumption. Hung (2005) goes further with credit rationing and its impaction on social consumption; he believes that credits for both investment and consumer loans are diminishing as capital accumulates, but together they are increasing as the government imposes more economic repression policies. Malkoc, Zauberman, and Ulu (2005) see credit as a tool to establish temporal distance and explore the impact on consuming decisions. Abdul Muhmin (2008) puts effort into examining the nature of consumer attitudes towards debt in Saudi Arabia. Banerjee, Ghosh, and Roy (2010) go further with the credit loan across rural households, exploring the attitude towards a loan, the network of lending and the credit constraint on the rural households. Brown and Woodruffe-Burton (2015) explore contemporary consumers' emotions and irrationality in attitudes towards indebtedness, using the UK payday loan data; from their research, it can be found that consumers will go beyond rational emotion, and payday loans will stimulate consumption. So, we can see that in these studies, credit constraint plays an important role on consumption, but there is still a lack of research that explores the relationship between credit constraint and consumption in China.

In this research, we focus on the consuming behavior of households faced with credit constraints. Using data from the 2011 China Household Finance Survey, we explore the role of credit constraint which may have an impact on consumer behavior. Because household credit constraints are not effectively quantified, we explore 2SLS regression to minimize the measure mistakes.

2.2.1 Method

We explored the impact of credit constraints on household consumption behavior; we selected the total consumption in the last year as the dependent variable, we and considered that the endogeneity of credit constraints would affect the accuracy of the empirical results. The instrumental variable (IV) was used for estimation. The method adopted is called 2SLS. The steps of the model were as follows.

Firstly, we modeled the question in a linear regression model, which is also called a structural equation, as (2.3). It is the first-stage model.

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + \beta_3 z_2 + \varepsilon \tag{2.3}$$

In the equation (2.3), the dependent variable y_1 represents the household consumption behavior (total consuming amount), y_2 is the endogenous variable, here y_2 is also on behalf of the credit constraint (self-reported constraint), and z_1 and z_2 are exogenous variables—here we chose house value and elderly people. House market value can represent one's wealth to some degree, and through the correlation test we found that house market value and consuming have no significant relation. Similarly, elderly people can represent one's family burden to some extent, and through the correlation test we found elderly people and consuming have no significant relation. Here we will not present the correlation test results. And then we estimated the coefficients of related variables using (2.3).

In the second stage, we built the ordinary least squares (OLS) regression using the estimate of the first stage. The specific regression model is as follows (2.4).

$$y_1 = \gamma_0 + \gamma_1 \hat{y}_2 + \gamma_2 z_1 + \gamma_3 z_2 + \varepsilon \qquad (2.4)$$

Because $\hat{y}_2$ is used instead of y_2, and $\hat{y}_2$ here is the instrument variable, which is borrowing amount, the 2SLS estimation is essentially different from the OLS estimation, which can reduce some measurement mistakes to a degree.

When constructing a model using the 2SLS estimation method, the instrument variable needs to satisfy two conditions. First, it must be irrelevant to the error term. Secondly, it must be strongly correlated with the explanatory variable. Corresponding tests on instrumental variables include weak instrument variable tests and over-identification tests. In addition, when the explanatory variable is exogenous, the validity of 2SLS estimation is not as good as OLS. At this time, the 2SLS estimator has a very large standard error, so it is necessary to perform an endogeneity test on the explanatory variable to verify whether the 2SLS estimation is better than OLS estimates.

The weak instrument variable test is mainly used to test whether the instrumental variable is correlated with the dependent variable. The weak instrument variable test can be found in the first stage of the 2SLS estimation model in Table 2-5.

Table 2-5 shows that individual instrumental variable—credit constraint—is strongly related to endogenous variables. In addition, in the 2SLS estimation method, the first-stage F statistic is greater than 10, which can be considered as a good result for a weak instrumental variable test, which is to say there is no weak instrument variable problem, or the instrument variable is good enough to fit this problem.

Table 2-5 Weakness Test Results of independent variables

Independent variable	Coefficient	P
self-reported constraint	-0.017***	0.006
house value	-0.001	0.187
elderly people	0.006	0.153
education	0.029	0.144
constant	0.076	0.172

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

The over-identification test is used to verify whether the instrumental variable is exogenous or not. Usually, it is done using the relationship with the error term. If the instrumental variables are endogenous, then the estimation method should be replaced or other instrumental variables should be used for estimation. With a single instrument variable, one cannot verify whether it is uncorrelated with the error term because the error term is unobservable. However, if there is more than one instrument variable, one can check whether the instrument variables are related to the error term or not by replacing the instrument variables. Statistically, the over-identification test is generally performed using a Sargan statistic. The Sargan statistic with self-reported constraints is 1.99 through replacement with borrowing, and the P value is 15.84%, greater than 15%. That is to say, at the significance level of 15%, there is no over-identification problem with the instrument variable—self-reported constraint.

The endogeneity test is used to test the endogeneity of explanatory variables. Here, we test endogeneity using the Hausman test. The basic principle of the Hausman test is to determine whether there is a systematic difference between the 2SLS estimation and the OLS estimation. The null hypothesis is that there is no systematic difference, and through the Hausman test the P value is 1.51%, less than 5%, indicating that the null hypothesis is rejected at a significance level of 5%. That is to say, there is a systematic difference between the 2SLS estimation method and the OLS estimation. Furthermore, it can be seen that the explanatory variables are endogenous (self-reported constraint is related to consuming), so the 2SLS estimation is more effective than the OLS estimation.

2.2.2 Empirical result and discussion

In this part, we use the 2SLS model to explore the relationship between consuming and household credit constraints. The specific regression results are presented in Table 2-6.

Table 2-6 Regression Results for consuming behavior

	2SLS		OLS	
Independent variable	Coefficient	P	Coefficient	P
self-reported constraint	-192.320**	0.012	354.286	0.231
house value	143.610	0.078	185.621***	0.003
elderly people	-69.190	0.053	-62.852***	0.002
constant	-59.240	0.144	-79.683	0.054
adjust R square	0.076		0.054	

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

Table 2-6 shows the 2SLS and OLS regression results of the impact of credit constraints on household consumption behavior.

From Table 2-6, it can be seen that whether using 2SLS or OLS estimation, house value has an obviously significant positive impact on household consumption at the significance level of 5%. Real estate accounts for a large proportion of the wealth of Chinese households. Especially in first-tier and second-tier cities, housing prices are rising rapidly. A large number of households which take part in the survey have no real estate, so they are likely to reduce their expenses and make more savings in order to buy a house. Meanwhile, the household will alleviate economic uncertainty; thereby, they will prefer to consume.

Furthermore, it can be found that the number of elderly people has negative relationship with consumption. Elderly people may impose some economic burden on a household. Households may not dare to consume when they face a lot of economic pressure. They can only rely on current wealth to make consumption possible, which makes consumption affected by fluctuations in the number of elderly people.

Through the 2SLS estimation, it can be found that self-reported constraint has a significant negative impact on household consumption; however, it is positive when using the OLS estimation because we know there is endogeneity between “self-reported constraint”, “elderly people” and “house value”, and thus we think the result of 2SLS will be better. So, it can be inferred that credit constraints will have a negative impact on Chinese household consumption behavior to a certain extent. This is concordant with the findings of many scholars in the past. Credit constraints usually inhibit the current period of borrowing by residents and then increase future uncertainty; as a result, they inhibit current consumption behavior.

So, from these results, we can see that improving the domestic consumption can begin with break credit constraints. Improving the financing environment and alleviating the credit constraints has a profound impact on promoting household consumption. Improving the financing environment should start from two aspects: one is to reduce credit ration, and the other is to reduce the financing cost, which can alleviate the economic uncertainty for households.

Moreover, establishing and improving the social pension service system is another effective way to improve consuming. China now faces an aging society problem, elderly people in every household are becoming more and more prominent, and the burden on families is getting heavier. The empirical results of this research show that the more elderly people there are in one’s family, the more likely credit constraints are, meaning that the family’s burden will affect credit constraints to a certain extent, and then it will restrict consumption. So, to break this cycle, China still needs to improve the social pension system.

But there are still many problems that should be discussed as to consuming and the credit constraint problem. If there were more quantitative data about credit constraint, we could go further with explorations, such as formal credit constraints and informal credit constraints, or demand-based credit constraints and supply-based credit constrains. More data would provide a more detailed and reliable basis for alleviating credit constraints, and would allow researchers to conduct more reasonable analyses with credit constraints and consuming.

2.3 Investment behavior under credit constraint

There have been many studies which discussed the relationship between investment and credit constraint (see the research of Eisfeldt and Rampini 2007; and Aghion, Angeletos, Banerjee, and Manova 2010), but they both take the macro view to discuss the relationship, and in their research, it can be found that credit constraint restricts investment to some degree. Kaboski, Lipscomb, and Midrigan (2014) put focus on household credit constraint; they establish a model of households with multiple demand such as smooth shocks, financing investments and constraints which are limited credit and self-control issues, and they intend to study the nature of household financing constraints in developing countries, so they try to relax the impact of these constraints. From their research, they illustrated that a short-term increase in access to loans leads to very distinct investment behavior in the short run. Insler, Compton, and Schmitt (2016) examined borrowers' investment decisions in relation to their cognitive ability (measured by the Cognitive Reflection Test, the SAT and grade point average), personality traits (captured by the Myers-Briggs Type Indicator), and other demographic characteristics; thereafter, they sought to relax credit constraint to a degree. Khan and Rouillard (2018) highlight the role of household borrowing constraints in accounting for investment; they argue that households' borrowing constraints in an economy can generate credit shortage crunch of US residential investment dynamics. To sum up, the recent studies demonstrate that credit constraint has an important effect on investment, but there lacks research which illustrates the relationship between credit constraint and investment in China. Is the credit constraint harmful for investment or not? How to relax the credit constraint? There is still a need for a lot of studies to explore these questions.

In this research, we put the focus on the investment behavior of households who face credit constraints. Using the data from the 2011 China Household Finance Survey, we explored the role of credit constraint which may have an impact on consumer behavior. Because household investment can be various forms from finance product to real estate, we used Tobit regression analysis to conduct this research.

2.3.1 Variables

There are 15 variables which have a direct impact on financing. Credit constraints variables are highly correlated with these variables, and we put the variables' descriptions in Table 2-7.

Table 2-7 Description of financing variables

Variables	Questions	Values
8B	If you borrow 100,000 RMB with the maturity in one year, how hard do you feel this is? (1 = very hard, 5 = very easy)	1–5
8C	If you borrow from relatives or friends, do you pay interest? (1 = no pay, 6 = higher than loan from bank)	1–6
8D	If you borrow from relatives or friends, which way do you prefer? (1 refers to mortgage; 2 refers to guarantee; 3 refers to borrowing; 4 refers to oral agreement; 5 refers to other.)	1–5
8GA	Do you have knowledge of house loans? (1 = know a little, 5 = know a lot)	1–5
8GB	Do you have knowledge of car loans?	1–5
8GC	Do you have knowledge of education loans?	1–5
8GD	Do you have knowledge of venture loans?	1–5
8GE	Do you have knowledge of consuming loans?	1–5
8GF	Do you have knowledge of renovation loans?	1–5
8JA	What do you think of getting house loans from banks? (1 = very hard, 5 = very easy)	1–5
8JB	What do you think of getting car loans from banks?	1–5
8JC	What do you think of getting education loans from banks?	1–5
8JD	What do you think of getting venture loans from banks?	1–5
8JE	What do you think of getting consuming loans from banks?	1–5
8JF	What do you think of getting renovating loans from banks?	1–5

In previous research, when several variables were relevant to credit constraints, factor analysis was used to reduce the number of variables. Following this approach, we conducted exploratory factor analysis. The results are shown in Table 2-8.

Table 2-8 Factor loading of financing variables

Factors	Variance	Percentage	Accumulative
1	4.779	48.898	48.898
2	1.071	10.957	59.855
3	0.673	6.888	66.742
4	0.485	4.958	71.701
5	0.468	4.792	76.493
6	0.421	4.303	80.796
7	0.386	3.947	84.743
8	0.304	3.112	87.855
9	0.268	2.747	90.602
10	0.244	2.501	93.103
11	0.202	2.065	95.169
12	0.152	1.554	96.723
13	0.129	1.318	98.041
14	0.101	1.035	99.076
15	0.090	0.924	100.000

From Table 2-8, we find that the loading scores of three factors exceed 85%, which indicates that 85% of the information in the original 15 variables can be replaced by merely three factors, as is described in the factor matrix in Table 2-9. We made a loading chart in order to further study the composition of these factors, which is shown in Figure 2-3.

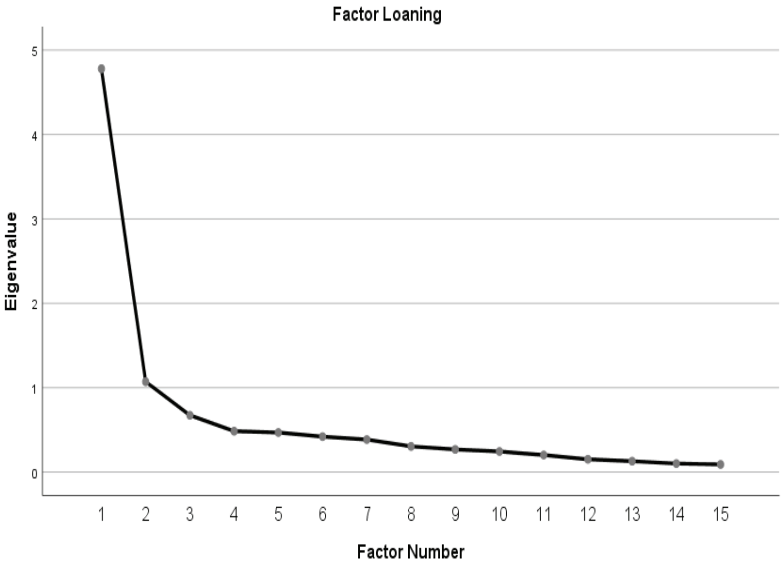

Fig 2-3 The eigenvalue loading of factors

Table 2-9 Transfer factor matrix of financing variables

	Original Variables			New Factors		
8B	-0.135	0.072	0.077	-0.341	0.182	0.194
8C	-0.163	0.010	0.094	-0.331	0.021	0.191
8D	0.901	-0.314	-0.047	0.788	-0.275	-0.041
8GA	0.930	-0.307	0.045	0.804	-0.265	0.039
8GB	0.816	-0.156	0.151	0.810	-0.155	0.150
8GC	0.583	0.008	0.203	0.629	0.008	0.219
8GD	0.710	0.309	0.032	0.706	0.307	0.032
8GE	0.782	0.367	-0.117	0.682	0.320	-0.102
8GF	0.670	0.366	-0.031	0.660	0.361	-0.031
8JA	-0.179	0.199	0.225	-0.330	0.366	0.414
8JB	-0.148	0.116	0.177	-0.352	0.275	0.419
8JC	-0.068	0.050	0.163	-0.162	0.119	0.390
8JD	-0.055	0.026	0.165	-0.102	0.049	0.308
8JE	-0.105	-0.061	0.186	-0.231	-0.134	0.411
8JF	-0.084	-0.030	0.231	-0.139	-0.050	0.383

2.3.2 Empirical result

To go one step further, we investigated the relationship between the three factors and investment behavior. A regression equation (2.5) was established. Its results are shown in Table 2-10.

$$y = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \beta_3 f_3 + \varepsilon \tag{2.5}$$

In the equation (2.5), the dependent variable y represents the household investment behavior (total investment amount), and f_1 , f_2 and f_3 are factor variables. We estimated the coefficients of the regression with OLS, and the result can be found in Table 2-10.

Table 2-10 Regression result of credit constraint

	Coef.	Std. Err.	P	[Lower Limit	Upper Limit]
f1	0.1749***	0.2046	0.0006	0.1640	0.1859
f2	0.4398***	0.1152	0.0021	0.3288	0.6108
f3	0.2812***	0.1089	0.0017	0.1942	0.3088
int.	11.0885	5.0732	0.1237	6.0795	15.0976

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 2-10, we can see that all the credit constraint factors have obviously positive impacts on investment behavior, which means that if a person is constrained by financing, they will have very little investment behavior.

Moreover, because there are many different types of investing behavior in our survey, we further divided the dependent variables into different ways of investing, including stocks, funds, debt, real estate and collection. And then we estimated the coefficients of related variables with multinomial Logit regression; the result can be found in Table 2-11.

Table 2-11 Multinomial Logit regression result of credit constraint

	Coef.	P	Coef.	P	Coef.	P	Coef.	P	Coef.	P
	stocks		funds		debt		real estate		collection	
f1	0.1853**	0.0294	0.1343**	0.0312	0.1266**	0.0394	0.8236***	0.0075	0.4853***	0.0094
f2	1.6482***	0.0012	1.5862***	0.0078	1.6482**	0.0214	1.2352***	0.0046	1.3222***	0.0062
f3	1.4452***	0.0024	1.5212***	0.0067	1.4452***	0.0089	1.2779***	0.0083	1.6452**	0.0483
age	0.0716	0.3491	0.4237	0.2134	0.0216	0.2618	0.0821	0.2442	0.0917	0.2855
male	0.1066	0.3142	0.1927	0.2749	0.0866	0.3382	0.0672	0.3161	0.0866	0.2437
marriage	0.3178	0.1775	0.1724	0.1276	-0.2178	0.1524	0.9226*	0.0932	-0.1248	0.1775
high school	-0.1724	0.3118	0.9456*	0.0925	0.1524	0.2132	0.1576	0.2116	0.0996	0.1622
undergraduate	0.2683	0.1667	0.7801	0.1328	0.2689	0.1124	0.1218	0.1572	0.1124	0.1734
income	0.9345	0.3862	0.4127*	0.0872	0.5623	0.2537	0.8943**	0.0453	0.0213*	0.0846
cash	0.9015	0.3963	0.5621*	0.0731	0.7035	0.2926	0.1243**	0.0352	0.0068	0.1122
intercept	-0.4127	0.3748	-0.1853*	0.0847	-0.8327	0.2814	-0.4326	0.2754	-0.2312	0.3518
Adjust R^2	0.3142									

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 2-11, we can see that credit constraints have a significant positive impact on all investment behavior; that is, the fewer credit constraints, the more investment behavior. At the same time, we can find that under the theoretical framework of the life cycle, investment behavior does not conform to the basic hypothesis of life cycle theory after controlling credit constraints. Age, gender and marital status have no significant impact on any investment behavior. Furthermore, education did not significantly affect investment behavior. The impact of income on investment is not as important as expected; only in the investment behavior in funds, real estate and collections is there a significant statistical impact. Among them, the impact on real estate investment is the most obvious. Every unit of income increase will have a positive impact of 0.89 lift on real estate investment. Similarly, the impact of cash on investment is mainly reflected in the investment in funds and real estate. The increase of cash will make consumers more inclined to invest in real estate.

In short, consumer credit constraints have a very important impact on consumer investment behavior. The more credit constraints consumers are subject to, the less investment consumers make. At the same time, because there is no uniform measurement of consumer credit constraints and they are closely related to other economic variables, when discussing the impact of credit constraints on consumer investment behavior, the result will inevitably have an endogenous problem. Therefore, we can only discuss the direction of the model, but the specific quantitative relationship needs to be further studied.

3. FINANCING DECISIONS

Financing decision-making is an important issue that every family needs to face. According to the life cycle theory of Ando and Modigliani (1963), consumers can solve the credit constraints in their life by financing, and then they can achieve the smooth effect on their personal wealth for their whole lifetime. Therefore, the core problem is how to analyze the financing behavior of consumers, how to make decisions when consumers are financing, and whether these financing decisions are rational or not.

In this chapter, firstly we discuss the basic framework of financing decision-making analysis, and following this framework we use the analytic hierarchy process to quantitatively analyze the decision-making process of household financing. Through the analysis, we can see two key factors: costs and convenience are the main factors which affect financing decisions. Then we discuss the financing decision difference among savvy groups. By distinguishing the differences in long-term borrowing and borrowing channels between families with financial backgrounds, we can see that appropriate financial education can help households to better use financial instruments and ease credit constraints. Finally, we describe the financing behavior of Chinese consumers through the actual survey data, and we conduct an empirical analysis from four factors to analyze the financing behavior. It can be seen that wealth, household structure and financial plan play an important role in long-term loan owning. And health status plays the key role in non-bank financing, while wealth restricts households' choices. The credit constraints have impacted the financing choices deeply.

3.1 Financing decision-making

The study of individual behavior decision originated in microeconomics decision process. There are two famous frameworks for individual decision-making; one is utility theory, the other is prospect theory.

In the study of the 1990s, researchers believe an individual makes a decision to maximize their utility, and then they put the focus on calculating utility. There are various heterogeneous frameworks, according to different decisions. Venkatesh, Morris, and Ackerman (2000) investigated gender differences in the overlooked context of individual adoption and sustained usage of technology in the workplace using the theory of planned behavior gender has various differences in the neglected context of individual decision and the technology use can be better explained by planned behavior theory. Verplanken and Holland (2002) examined the value-behavior relation and focused on motivational properties of values, the self and value activation. Vigil (2009) argued that risk evaluation is the key indicator in personal decisions. Auh and Shih (2005) attempted to build and extend the literature on financial decision-making by drawing on mental accounting, which follows prospect theory. Levin, Bossard, Gaeth, and Yan (2014) follow the prospect theory, and attempt to check different types of framing effects; they uncover several key differences across ages, including levels of risk taking and sensitivity to expected value differences between risky and riskless choices.

Nowadays, researchers have expanded the utility framework of decision-making, using a contingent-claim model that explicitly considers long-run consumer welfare under uncertainty. Related work can be found in the studies of Becker and Shabani (2010) and Cunha, Lambrecht, and Pawlina (2011). It can be inferred that, whether utility theory or prospect theory, the basic assumption is that the decision-maker can obtain the greatest benefit by choosing the necessary decision-making scheme. So, in our work, we put the focus on how to evaluate these decision-making schemes, and we intend to provide a quantitative decision-making method.

Thus, we follow the framework of the analytic hierarchy process (AHP) to perform qualitative analysis on the financing behavior, which can combine qualitative discrimination with quantitative decision-making, and stratify and quantify people's thinking processes. According to the nature and purpose of the problem, the AHP breaks down the utility criterion into indicators with different levels. Then we can compare each indicator at the same level, thereafter determine its relative importance, and then calculate the importance as weights; thus, combining the weights allows us to calculate the maximum utility of decisions. Through the process of maximum utility, we can explore the decision-making process following this framework. Furthermore, our survey questions were extracted from the ones used in the China Survey of Consumer Finance conducted by the Chinese Financial Center of Tsinghua University in 2011 and the data collected from the student families as a side survey for the third China Household Finance Survey in 2015.

3.1.1 AHP analysis of financing choices

The analytic hierarchy process was officially proposed by T. L. Saaty in the mid-1970s (Saaty 2003, 2013). It is a qualitative and quantitative, systematic and hierarchical analysis method. It has four steps, which are: establishing a hierarchical structure model, constructing a comparison matrix, calculating weight vectors and checking consistency. Following these steps, the method can put the qualitative decision process into quantitative weights.

In this study, we set our decision goal as obtaining the maximum consumers' utility. And we broke down consumers' financing utility into five quantitative indicators, which are: financing convenience, financing costs, financing quotas, financing speed and repayment requirements. Then we sent the questionnaires to consumers and let them rank the five indicators; for instance, a_{ij} means the relative position for the convenience indicator. The position is listed by importance, which is marked by the sequence of 1,3,5,7,9. $a_{ij} = 1a_{ij} = 1$ indicates that a_i and a_j are equally important; $a_{ij} = 9$ indicates that a_i is significantly more important than a_j. Then, from every questionnaire, we can get the evaluation matrix as:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}.$$

From the Chapter 1, we know there are four main financing channels in China, which are: private borrowing, non-bank institution, bank and others. Then the decision can be broken down as in Figure 3-1.

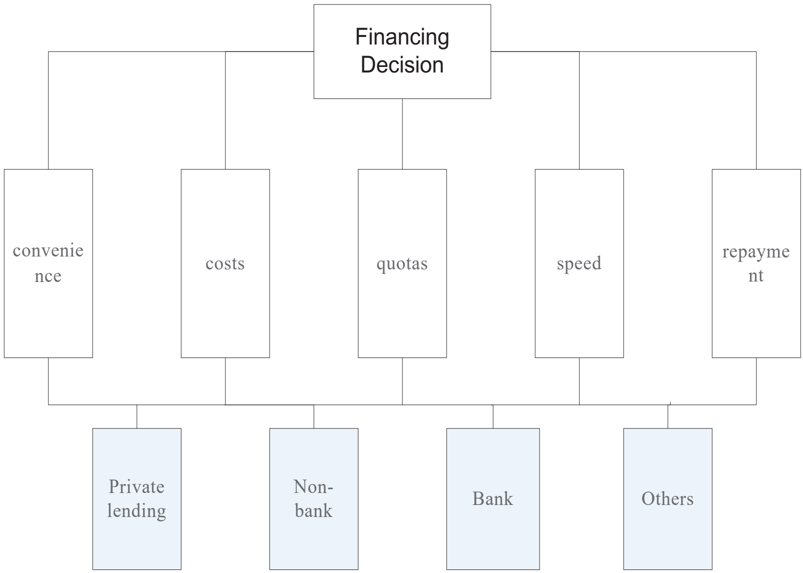

Fig 3-1 Graph of financing decision breakdown

The ranks of the criterion must satisfy the consistency check, which is shown in the equation (3.1)

$$\mathrm{CI} = \frac{\lambda_{\max}(A) - n}{n-1} \tag{3.1}$$

where n is the dimension of the matrix **A**, λ is the maximum eigenvalue of **A** and $U = (u_1, u_2, u_3, u_4, u_5)$ is the eigenvector of λ.

Meanwhile, if the questionnaires are conducted randomly, there will also be the consistency index (CI), so we should compare the CI with the random consistency index (RI). The random consistency is listed in Table 3-1. Thereafter, the random consistency indicator can be calculated as equation (3.2).

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{3.2}$$

When CR <0.1, the criterion matrix satisfies the consistency and is completely acceptable; otherwise, the matrix does not satisfy the consistency and needs to be discarded. If the criterion gets through the consistency check, we can apply it to make the decision.

Table 3-1. The index of RI

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

Then we select the four decisions through the same criterion, and the characteristic of decisions x can be calculated by rank matrixes through comparison with each other; for example, for the first financing decision, private borrowing can be evaluated by rank as in equation (3.3).

$$b_1 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \tag{3.3}$$

Through calculating all the eigenvalues of rank matrixes, we can get the maximum eigenvalue θ_x and the eigenvector w_{xj}. If this value can pass the random consistency test as in (3.2), then we can use the eigenvectors w_{xj} and u_j to quantify the decision as in equation (3.4).

$$W(x) = u_j w_{xj} \tag{3.4}$$

Finally, the weighted preference of the four financing decisions corresponding to each questionnaire can be obtained through calculation.

3.1.2 Statistics of financing choices

Through the whole calculation, we obtained a total of 623 valid questionnaires. And the statistics of the five criteria can be seen in Figure 3-2.

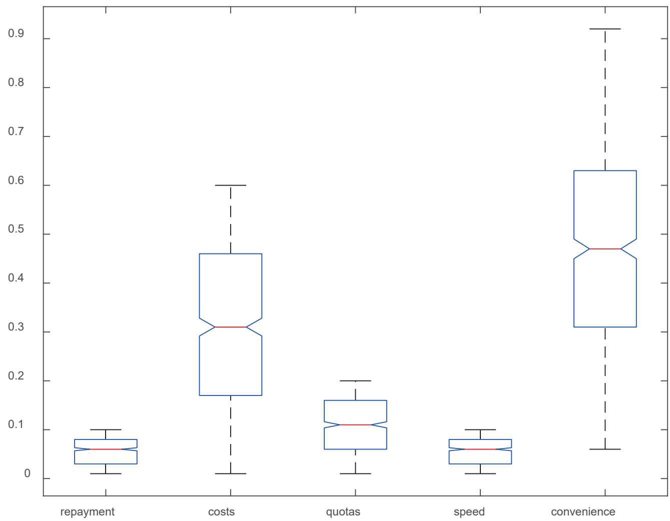

Fig 3-2 The distribution of the eigenvalues of the five criteria

As shown in Figure 3-2, we can find that cost and convenience are the two most important concerns for financing decisions, and there is more diversity with the distribution of convenience than other criteria.

For more explorations, we checked the final ranks of the four financing channels, which can be seen in Figure 3-3.

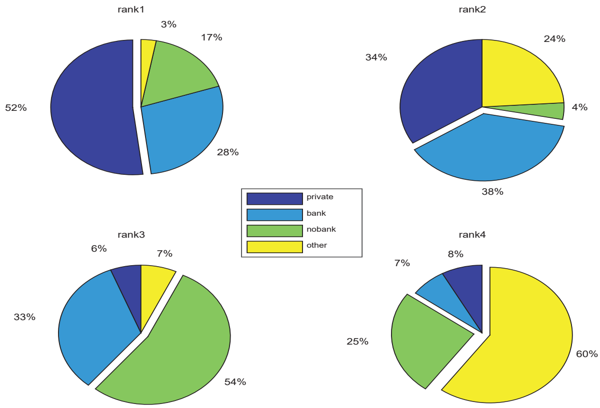

Fig 3-3 The ranks of the four financing channels

From Figure 3-3, it can be found that in most instances private is the first choice for consumer financing; and nearly half of persons will choose to borrow money from relatives and friends. Non-bank is probably the last choice for consumer financing; the rank rises sharply with the third and last choice.

To sum up, from the results of the analytic hierarchy process, we can find that most consumers agree that the cost of capital and the convenience of borrowing are the most important indicators when making financing decisions. Therefore, private borrowing is the first choice for household financing, and it will still exist for a long time in the future in China.

3.2 Decision-making difference among savvy groups

Decision-making processes may be greatly influenced by information. García (2013) argues that individuals who use information in making financial decisions reduce some psychological states such as overconfidence, and he aims to complement this emerging focus by studying how individuals perceive and process information when making financial decisions. Lusardi and Mitchell (2014) designed surveys to establish how much (or how little) people know and identify the least financially savvy population subgroups, and then they examine the impact of financial literacy on economic decision-making in the United States and elsewhere. Collins and Odders-White (2015) begin a financial education course among high school students; they emphasize "the need for understanding the underlying mechanisms that facilitate the translation of student knowledge into the ability to make sound economic decisions over the life". Cole, Paulson, and Shastry (2016) argue that financial literacy and cognitive capabilities are convincingly linked to the quality of financial decision-making. The work of Chenavaz (2017) introduces the time point into the decision-making process; through studying the dynamic quality policy when consumers use a reference point in their decision-making, they get in line with the principles of behavioral economics standards. Fan and Chatterjee (2017) claim that financial literacy, perceived financial knowledge, educational attainment and engaging the services of a financial professional are positively associated with the likelihood of financing decision-making process, which are the key factors that affect consumers' behavior. Kim, Gutter, and Spangler (2017) emphasize that financial professionals and educators need to integrate family members into financial education and counseling.

In total, it can be found that information plays an important role in consumer behavior, and the professional financial intellectuals will behave different from others. So we will put our focus on the difference in financial decisions among different loan holder groups. Through comparisons among savvy groups, we intend to infer the special characteristics of financial intellectuals. We divided this part into three subsections: one is about the long-term loan behavior; another is about financing choice; and the final one is about financial plans.

3.2.1 Long-term loan behavior

We define the "long-term loan" as liabilities with maturity over one year or more. Based on the discussion above, we divided the study participants into groups by region, education level, industry, population and health status, and we observed the long-term loan behavior in different groups.

According to the National Bureau of Statistics, the cities can be divided into three categories, which are: first-tier city, second-tier city and third-tier city; the results can be seen in Table 3-2. The value ratio means the total loan value percentage among different regions, and the holding rate means the percentage of holders among different regions. It can be seen from Table 3-2 that house loans make up the majority of total long-term loans, and nearly 20% of households in first-tier cities have mortgages.

Table 3-2. The long-term behavior among different regions

		tier 1	tier 2	tier 3
house loan	mean	54,467	22,266	16,703
	value ratio	55.58%	10.71%	33.71%
	holding rate	19.71%	15.50%	13.44%
car loan	mean	2,453	983	1,027
	value ratio	49.58%	9.36%	41.06%
	holding rate	2.54%	1.84%	2.23%
other long-term loan	mean	3,007	793	1,932
	value ratio	41.75%	5.19%	53.06%
	holding rate	2.48%	2.43%	2.61%

Long-term financing at different educational levels is shown in Table 3-3. From Table 3-3, it can be seen that the higher the education level is, the more house loans there are. But doctors rarely have car loans and other long-term loans. And for the value ratios, bachelors have the largest value ratios, so it can be inferred that the bachelors have larger single long-term loan amounts than others.

Long-term loans with different industries are shown in Table 3-4. Table 3-4 shows that the household members who work in different industries have different long-term loans. As can be seen, households with members who work in the public sector have the most monetary amount of house loans from banks; households with members who work in foreign industries have the most monetary amounts of house loans from relatives and car loans. For the holding rate, there seems to be no significant difference among households with house loans; furthermore, it can be seen that the house loan's holding rate is about 20%, which is more than others.

Table 3-3. The long-term behavior among different education levels

		high school	bachelor	master	doctor
house loan	mean	21,343	41,879	72,231	144,639
	value ratio	24.98%	61.31%	6.06%	7.76%
	holding rate	17.21%	22.37%	28.15%	58.19%
car loan	mean	852	2,509	4,366	3,302
	value ratio	19.70%	72.61%	7.26%	0.43%
	holding rate	5.00%	5.82%	6.77%	1.74%
other long-term loan	mean	1,561	3,143	448	0
	value ratio	24.88%	62.67%	12.45%	0%
	holding rate	5.98%	5.09%	1.49%	0.00%

Table 3-4. The long-term behavior among different industry

		Public	Private	Foreign	Others
house loan from bank	mean	40,216	37,139	29,150	14,198
	value ratio	49.19%	33.36%	11.20%	6.24%
	holding rate	21.57%	22.90%	18.32%	14.60%
house loan from relatives	mean	14,848	8,366	10,858	11,210
	value ratio	16.37%	12.16%	60.07%	11.39%
	holding rate	22.57%	20.21%	18.56%	26.41%
car loan	mean	1,898	1,312	2,120	794
	value ratio	17.84%	16.26%	59.02%	6.88%
	holding rate	5.31%	5.76%	3.99%	4.30%
other long-term loan	mean	5,990	794	2,006	517
	value ratio	38.57%	6.74%	51.62%	3.07%
	holding rate	5.58%	6.24%	4.08%	2.49%

Long-term loans among households with different structures are shown in Table 3-5, Table 3-6 and Table 3-7. Table 3-5 shows the differences in long-term loans between all different types of households. Table 3-6 shows the differences in long-term loans between households with different numbers of children, and Table 3-7 shows the differences in long-term loans between households with different numbers of elderly people.

As can be seen from Table 3-5, long-term housing loans are the main part of long-term household loans. Most households have more than 20,000 RMB of long-term loans from banks; long-term loans from relatives increase with the number of household members. About 4% of households have long-term car loans, and there is no significant difference in the number of households with long-term car loans.

Table 3-5. The long-term behavior among households of different sizes

		1	2	3	4	>=5
house loan from bank	mean	25,275	32,095	23,856	31,451	31,312
	value ratio	26.09%	15.42%	28.35%	9.53%	20.61%
	holding rate	12.68%	18.47%	15.99%	20.26%	18.12%
house loan from relatives	mean	3,088	7,740	9,445	14,028	20,423
	value ratio	8.90%	10.38%	31.33%	11.86%	37.52%
	holding rate	8.21%	15.41%	21.95%	24.45%	32.36%
car loan	mean	619	1,684	1,321	1,640	1,389
	value ratio	14.42%	18.27%	35.45%	11.21%	20.64%
	holding rate	2.47%	4.48%	5.55%	4.66%	4.52%
other long-term loan	mean	1,088	2,135	1,770	2,594	2,098
	value ratio	17.49%	15.98%	32.78%	12.24%	21.51%
	holding rate	3.27%	6.71%	4.94%	5.30%	4.68%

Table 3-6. The long-term behavior among households with different numbers of children

		none	1	2	>=3
house loan from bank	mean	33,338	25,765	29,672	8,127
	value ratio	42.68%	44.12%	12.27%	0.92%
	holding rate	17.27%	19.48%	16.16%	2.36%
house loan from relatives	mean	8,205	13,702	17,648	14,634
	value ratio	24.47%	54.66%	17.00%	3.86%
	holding rate	16.67%	26.40%	27.19%	31.13%
car loan	mean	1,299	1,558	1,739	0
	value ratio	32.93%	52.83%	14.24%	0.00%
	holding rate	3.71%	6.26%	3.66%	0.00%
other long-term loan	mean	1,982	2,160	1,004	6,165
	value ratio	34.52%	50.33%	5.65%	9.50%
	holding rate	4.95%	5.38%	3.89%	7.50%

As can be seen from Table 3-6, households without children have the largest monetary amount of long-term loans from banks; households with one child mainly have long-term loans from relatives and friends. But interestingly, as for the holding rate, when the number of children is more than or equal to three, few families can take out long-term loans from banks; most of the long-term loans come from relatives and friends. For the long-term car loans, it can be seen that the holding rate is relatively low, and the value is relatively small. Similarly, when there were more than three children in the household, there was no long-term car loan in the households who were surveyed.

Table 3-7. The long-term behavior among households with different numbers of elderly people

		none	1	2	3	>=4
house loan from bank	mean	14,240	16,774	27,007	44,657	54,139
	value ratio	14.13%	7.75%	30.86%	13.00%	34.26%
	holding rate	11.54%	14.42%	17.19%	25.95%	27.29%
house loan from relatives	mean	7,868	10,182	14,679	18,343	13,152
	value ratio	18.18%	10.95%	39.05%	12.44%	19.38%
	holding rate	18.06%	17.52%	25.27%	30.46%	27.93%
car loan	mean	725	1,853	1,236	1,412	2,613
	value ratio	14.25%	16.95%	27.95%	8.14%	32.72%
	holding rate	3.95%	5.80%	4.77%	5.86%	5.23%
other long-term loan	mean	871	1,522	2,176	2,461	4,083
	value ratio	11.75%	9.56%	33.81%	9.75%	35.13%
	holding rate	3.64%	4.70%	6.28%	5.85%	5.31%

As can be seen from Table 3-7, in terms of monetary value ratio, households that need to support four or more elderly people have the most long-term loans from banks, households that need to support two elderly people have the most long-term loans from relatives and friends, and households that need to support four or more elderly people have the most amount of car loans. In terms of holding ratios, it seems that the higher the number of elderly people whom the household needs to support, the higher the percentage of households with long-term loans is.

Therefore, we can conclude that the decision-making of long-term loan behavior, on the one hand, is affected by the household's own background and information; on the other hand, it is also affected by the financing attitude to the policy. Generally speaking, policy can change the long-term financing attitude, and most households are cautious about long-term loans.

3.2.2 Financing choice preference

In the foregoing, we have discussed the differences in households' long-term loan behavior. Family structure and education level largely determine the family's risk preference. To a certain extent, regions and industries influence the credit policy; thus, risk preference may change household decision-making. In this part, we focus on the household's attitude towards different attributes of financing instruments under different risk attitudes, so as to analyze the reasons for their preferences of financing instruments in subjective decision-making.

The following question is designed to investigate consumers' risk attitudes quantitatively labeled as risk aversion: "Assume that a coin is tossed; you will get 2,000 RMB if it comes up heads, but you will get nothing if it comes up tails. Supposing you resell such a profit opportunity, what is the minimum amount you would charge for it?" w_i represents the initial wealth of the respondent i, $u_i(\cdot)$represents the utility function of i and x_i represents the lowest bid of the respondent in the game. Then the respondent's utility equation in this game can be expressed as equation (3.5),

$$u_i(w_i + x_i) = 0.5u_i(w_i) + 0.5u_i(w_i + 2000) = E[u_i(w_i + x_i + P_i)] \quad (3.5)$$

where $E(\cdot)$ represents mathematical expectation and P_i represents random income in the game. Using second-order Taylor expands（3.6） at $w_i + x_i$:

$$u_i(w_i + x_i) \approx u_i(w_i + x_i) + u'_i(w_i + x_i)E[P_i] + 0.5u''_i(w_i + x_i)E[{P_i}^2] \quad (3.6)$$

The respondent's absolute risk aversion, which we call as "Ara", can be expressed as follows after simplification:

$$Ara_i(w_i + x_i) = -\frac{u''_i(w_i+x_i)}{u'_i(w_i+x_i)} = \frac{2E[P_i]}{E[{P_i}^2]} = \frac{2000-2x_i}{2000000+x_i^2-2000x_i} \quad (3.7)$$

For more detail, see the work of Han, Xiao, and Su (2018).

The distribution of Ara in each sample is shown in Figure 3-4. The original distribution is shown in the right part of Figure 3-4. We changed the interval so that the cumulative probability of the whole sample distribution in each interval is approximately consistent with the normal distribution. The continuous variable is divided into five groups according to their distributions, in which [-1,-0.8] = risk aversion and is assigned 1; [-0.8,-0.2] = slight risk aversion and is assigned 2; [-0.2,0.2] = risk neutrality and is assigned 3; [0.2,0.6] = slight risk preference and is assigned 4; and [0.6,1] = risk preference and is assigned 5.

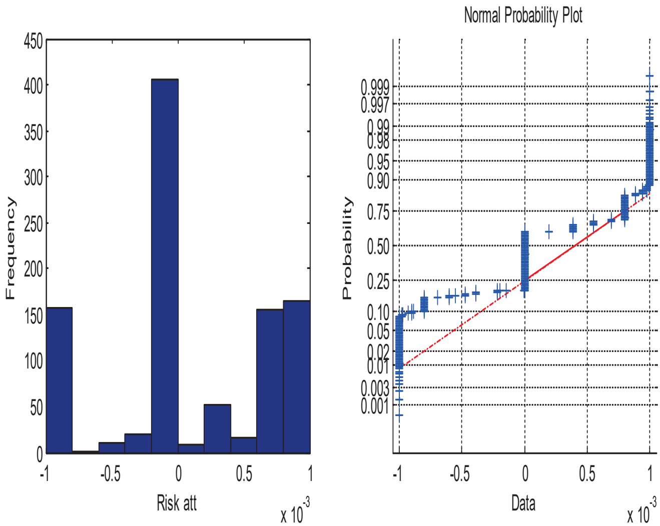

Fig 3-4 The Distribution of Risk Attitude

The first financing choices under different risk attitudes are shown in Table 3-8. As can be seen from Table 3-8, most households choose to borrow from relatives and banks. With the increase of risk preference, the proportions of non-bank lending and other lending have increased. At the same time, it can be noted that more than 40% of people will take bank lending as the preferred financing method, which shows that Chinese households have begun to have a clearer understanding of financing channels. Interestingly, for non-bank lending, the degree of risk preference and financing choice cannot directly reflect the trend between the relationships. It can be seen that the lowest proportion is about 2.36%, corresponding to the group with lower risk preference as 2, and the second-lowest proportion is about 3.65%, corresponding to the group with higher risk preference as 4. Because of the data, it is difficult to explain this phenomenon here.

Table 3-8. First financing choice under different risk attitudes

	1	2	3	4	5
Private	51.45%	56.71%	42.15%	43.21%	36.84%
Bank	43.22%	39.16%	44.23%	51.26%	48.26%
Non-Bank	5.17%	2.36%	9.74%	3.65%	10.28%
Other	0.16%	1.77%	3.88%	1.88%	4.62%

Similarly, the relationship between the last financing options and risk preferences is analyzed as shown in Table 3-9. From Table 3-9, we can see that there is no obvious relationship between the last financing choice and risk preference. At the same time, the proportion of private loans, bank loans and non-bank loans are all relatively small. Other loans account for the largest proportion, more than 30%. Thus, it is impossible to infer this behavior directly from the data, but it can be simply inferred that the last financing choice cannot be used as an effective variable for analysis.

Table 3-9. Last financing choice under different risk attitudes

	1	2	3	4	5
Private	13.23%	14.24%	13.88%	11.52%	10.26%
Bank	9.61%	7.63%	8.52%	9.37%	10.66%
Non-Bank	7.26%	13.82%	8.55%	9.61%	9.77%
Other	32.54%	38.88%	33.25%	42.19%	36.21%

3.3 Factors influencing financing behavior

Since the 1950s, developed countries have formulated financial policies to regulate macro-consumption and stimulated consumption by expanding lending. With the development of the economy, loans have become an important part of consumption channels. According to the China Survey of Consumer Finance conducted by the Chinese Financial Center of Tsinghua University in 2011, the participation rate of household borrowing activities in China is as high as 33.5% of total households. Meanwhile, no more than 30% of these households can get loans from banks. This reflects that the supply level of financial services in China's banks cannot match the demand of households for borrowing, such that many households take loans through private borrowing channels. And in recent years, the number of illegal cases caused by private borrowing experienced exponential growth, which has caused serious social problems. Therefore, an in-depth analysis of the current situation of household financing behavior and influencing factors can not only provide a relevant basis theory for household financial planning but can also have important significance to making relevant policies to protect consumers in China.

According to Han and Li (2011), the factors influencing household financing behavior can be divided into two categories: one is macroeconomic variables—i.e., current interest rate, labor participation rate, inflation rate, credit aggregates, and so on; the other is micro variables—i.e., household education, household region, family size, household assets, income level, investment and consumption characteristics, and so on.

In this research, we mainly put our focus on the impact of micro characteristics on household borrowing behavior and the discrimination of household borrowing preference—that is, the preferred channel of borrowing.

Based on the above analysis of financing decision-making, we put our focus on the long-term loan and the financing channels. As for the long-term loan ratio, it is a continuous variable, and we have checked the distribution, if it can be divided into four groups, which are: below 20%, [20%,50%], [50%,80%] and above 80%. And for financing channels, we chose four financing channels as dependent variables: 1 refers to private borrowing; 2 refers to bank borrowing; 3 refers to non-bank institute borrowing; 4 refers to others.

3.3.1 Variables

Following the research of Huston (2012); Ruiz Tagle and Vella (2016) and Han, Xiao, and Su (2018), we chose 17 micro variables for this empirical research, and these independent variables can be divided into four categories. First, wealth indicators include region, total assets, annual income and cash income proportion. Second, the structure of the household includes total household population, number of children, number of workers and number of elderly people. Third, the household's personal background includes marriage, gender, region, education, working years, health status and industry. Last, financial plan includes financial planning and risk preference. The details of the independent variables can be found in Table 3-10.

Among these variables, cash proportion is defined as the cash or instable income percentage of the total income, and it is a continuous variable, but its distribution does not follow normal distribution, so we divided it into four categories. The first is below 10% which is referred to as 1; the second is [10%,20%] which is referred to as 2; the third is [20%,30%] which is referred to as 3; the last is above 30% which is referred to as 4.

We used the data from the China Survey of Consumer Finance conducted by the Chinese Financial Center of Tsinghua University in 2011. There were 5,911 households; we omitted the samples without answers for the questions and the samples that have no financing behavior, and there are still 3,060 samples in our research. The statistical descriptions of some variables are shown in Table 3-11 (a) and (b).

From Table 3-11 (a), it can be found that in this sample, most households come from the second-tier cities females are the majority, more than a half of these family members have a college degree or higher education experience, most of these household members work in public industry and the working experience is less than 10 years. And it can be found that almost all of these households have financial planning. It is, as usual, the case that most of these households are risk averse. From Table 3-11 (b), we can find that the mean income of households is 50,310 RMB, and the mean asset level of households is 732,470 RMB. We transferred these two variables into normal distribution.

Then we took the wealth variables into a cross relationship validation test; the results can be found in Table 3-12. From the Table 3-12, in terms of household income, the incidence of households with annual income less than 50,000 RMB is 36.2%, and the preferred borrowing channel is private borrowing with the ratio of 79.2%; the incidence of households with annual income between 50,000 RMB and 200,000 RMB is 73.7%, and the preferred borrowing channel is private borrowing with the ratio of 74.5%; the incidence of households with annual income more than 200,000 RMB is 42.8%, and the preferred borrowing channel is the same as the others'. So, it can be inferred that households in the middle class are more likely to have a long-term loan. In terms of total household assets, the incidence of long-term loan owner percentage at medium asset level is significantly higher than that of the other two categories. And for the financing channels, all the households prefer private borrowing; the fewer assets a household owns, the more likely they will borrow from non-bank institutions. The incidence of borrowing behavior corresponding to annual income and total assets is quite different. From only the simple descriptive statistics, we cannot summarize the impact of household income and assets on the financing behavior.

Next, we took family structure into a cross relationship validation test; the results can be found in Table 3-13. From Table 3-13, it can be inferred that regardless of family structure, nearly half of households have long-term borrowing behavior, and most households use private borrowing as the primary means of financing.

Table 3-10 The Summary of Independent Variables

Class	Variables	Type	Note
Wealth	total asset	continuous variable	Conversion to Standard Normal Distribution.
	annual income	continuous variable	Conversion to Standard Normal Distribution.
	cash proportion	continuous variable	
Structure	population	discrete variable	
	children	discrete variable	
	workers	discrete variable	
	elderly people	discrete variable	

Background	region	discrete variable	According to the National Bureau of Statistics, cities are divided into categories one, two and three.
	marriage	binary variable	
	gender	binary variable	
	education	discrete variable	Junior high school and below = 1; high school and vocational school = 2; three-year or four-year college = 3; master’s degree = 4; doctoral degree = 5.
	industry	discrete variable	Public industry = 1; private industry = 2; foreign industry = 3.
	working years	discrete variable	0–5 = 1; 5–10 = 2; 10–20 = 3; 20–30 = 4; more than 30 = 5.
	health status	discrete variable	No one was hospitalized in the past three years = 1; no one was hospitalized in the past year = 2; other =3.
Financial plan	financial planning	binary variable	Has financial planning = 1; other = 0.
	risk preference	discrete variable	1 = will not accept any risk; 2 = can accept some risk; 3 = risk neural; 4 = can accept much risk; 5 = can accept any risk.

Table 3-11 Statistical description of variables (a)

Variables	Class	Frequency	Percentage
region	first-tier city	916	0.2993
	second-tier city	1,471	0.4807
	third-tier city	673	0.2199
marriage	married	1,842	0.6020
	unmarried	1,218	0.3980
gender	male	968	0.3163
	female	2,092	0.6837
education	1. below high school	124	0.0405
	2. high school	336	0.1098
	3. undergraduate	1,312	0.4288
	4. graduate	1,126	0.3680
	5. doctor	162	0.0529
industry	1. government	1,457	0.4761
	2. public company	1,128	0.3686
	3. private company	475	0.1552
working years	1	873	0.2853
	2	1,247	0.4075
	3	654	0.2137
	4	241	0.0788
	5	45	0.0147
health status	1	1,763	0.5761
	2	2,867	0.9369
	3	673	0.2199
cash proportion	below 10%	1,327	0.4337
	10%–20%	211	0.06890
	20%–30%	122	0.03987
	above 30%	1,400	0.4575
financial planning	0	183	0.0598
	1	2,877	0.9402
risk preference	1	931	0.3042
	2	1,622	0.5301
	3	344	0.1124
	4	88	0.0288
	5	75	0.0245

Table 3-11 Statistical description of variables (b)

Variables	Mean	Standard Deviation	Min	Max
income (thousand yuan)	50.31	12.29	1.01	1,234.66
assets (thousand yuan)	732.47	619.26	4.26	10,832.83

Table 3-12 Cross relationship between wealth and financing behavior

Variables	Class	Long-term loan	Private borrowing	Bank borrowing	Non-bank borrowing
Income (thousand yuan)	<50	36.20%	79.20%	11.80%	0.31%
	50–100	73.21%	74.50%	22.10%	1.32%
	>100	42.68%	64.40%	30.26%	5.33%
Assets (thousand yuan)	<500	31.20%	41.22%	15.62%	18.23%
	500–1,000	42.33%	78.82%	18.42%	0.53%
	>1,000	28.68%	71.24%	23.22%	4.16%
Liability (thousand yuan)	<100	31.20%	41.22%	15.62%	18.23%
	100–500	42.33%	78.82%	18.42%	0.53%
	>500	28.68%	71.24%	23.22%	4.16%

Table 3-13 Cross relationship between family structure and financing behavior

Variables	Class	Long-term loan	Private borrowing	Bank borrowing	Non-bank borrowing
population	<3	39.7	62.11	23.15	8.37
	3	32.6	74.12	15.12	3.28
	>3	43.1	54.23	28.72	4.56
children	0	39.8	44.23	46.88	3.24
	1	39.1	53.27	41.18	4.22
	2	38.2	51.26	40.19	4.26
	>2	41.3	62.12	34.02	3.28
workers	<2	34.7	62.18	27.14	4.24
	2	41.1	55.27	31.16	4.67
	>2	33.9	52.63	34.88	3.35
elderly people	<2	29.7	59.23	29.25	8.34
	2	36.6	62.14	25.36	6.72
	3	49.2	64.33	24.26	5.83
	4	49.3	72.45	21.18	5.36
	>4	33.3	77.35	15.42	3.25

3.3.2 Empirical study

In this section, we separated our question into two parts. The first one is the factors which impact the loans that a household has. If a household borrows in the form of long-term loans, we marked the dependent variable as 1; otherwise as 0. The second part is the factors which impact the financing channels; we marked the four channels as follows: private borrowing as 1; bank borrowing as 2; non-bank borrowing as 3; others as 4.

Firstly, we used the Spearman correlation coefficient to examine the relationship between long-term loans and 16 micro variables. We omitted total liability from the examination because long-term loans can be seen as part of total liability. The relationship can be found in Table 3- 14.

Table 3-14 Spearman correlation for long-term loan borrowing

	Spearman correlation coefficient	Significance
total assets	0.031*	0.089
annual income	0.054**	0.030
cash proportion	0.040**	0.025
population	0.069***	0.000
children	-0.027	0.134
workers	-0.005	0.780
elderly people	-0.111***	0.000
region	0.095***	0.000
marriage	-0.027**	0.008
gender	0.032	0.211
education	0.118***	0.000
industry	-0.085**	0.004
working years	0.040**	0.025
health status	0.059***	0.001
financial planning	0.326***	0.000
risk preference	0.317***	0.000

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 3-14, it can be inferred that total assets, annual income, cash proportion, population, number of elderly people, region, marriage, education, industry, working years, health status, financial planning and risk preference have relationships with long-term loan borrowing. Then we used logistic regression to check the relationship between these factors and having a long-term loan; the result can be found in Table 3-15.

Table 3-15 Logistic regression for long-term loan borrowing

Long-term loan borrowing	Coef.	P
total assets	0.124**	0.029
annual income	0.163***	0.001
cash proportion	0.184***	0.002
population	0.122	0.349
elderly people	0.107	0.314
region	-0.318	0.178
marriage	0.172	0.312
education	0.268	0.167
industry	0.037	0.386
working years	0.084**	0.022
health status	-0.413**	0.037
financial planning	0.183**	0.027
risk preference	0.142**	0.032
adjust R square	0.242	

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 3-15, we can infer that only total assets, annual income and cash proportion have positive correlations with long-term loan borrowing, so it can be said that if household has more wealth, it will be more likely to borrow in the form of a long-term loan. The family structure has no significant relationship with long-term loans. For the background variables, there are only two variables have relationships with long-term loans: Working years has a positive relationship with long-term loans, so it can be inferred that if one has more working years, they will more likely to have a long-term loan. Moreover, health status has a negative relationship with long-term loans; this means that if the household has a good health condition, it will be less likely to have a long-term loan. For the financing plan variables, it can be found that financial planning and risk preference both have positive relationships with long-term loans; this means that if a household has a financial plan, and if the members are risk preferring, they will be more likely to have a long-term loan.

Next, we went further with the financing channels. Because there are four choices for financing channels, we chose multinomial logistic regression. Multinomial logistic regression is generally used to analyze the influencing factors when there are multiple options for dependent variables (equivalently categorical, meaning that options falls into any one of a set of categories that cannot be ordered in any meaningful way). Here we assume the independence of irrelevant alternatives for the four choices because, from the section of financing decision-making, we can see the financing order is fixed by household, so financing order choice will not change by other alternatives. The result can be found in Table 3-16.

Table 3-16 Multinomial logistic regression for financing choice

	Private Borrowing		Bank Borrowing		Non-bank Borrowing		Others	
	Coef.	P	Coef.	P	Coef.	P	Coef.	P
total assets	0.054	0.035**	0.247	0.004***	-0.074	0.058*	0.013	0.182
annual income	0.133	0.174	0.253	0.003***	-0.022	0.073*	0.012	0.164
cash proportion	-0.082	0.166	0.114	0.218	-0.082	0.046**	-0.005	0.078***
population	0.036	0.253	0.022	0.237	0.083	0.162	0.017	0.114
children	-0.142	0.086*	0.043	0.354	-0.054	0.128	0.011	0.163
workers	-0.008	0.144	0.018	0.277	-0.089	0.132	0.126	0.079***
elderly people	0.124	0.072*	0.112	0.183	-0.032	0.087*	0.113	0.119
region	0.076	0.041**	0.005	0.452	-0.017	0.095*	0.142	0.231
marriage	-0.053	0.337	0.032	0.164	0.022	0.211	-0.083	0.266
gender	0.046	0.159	0.011	0.196	0.018	0.185	0.052	0.254
education	0.091	0.097*	0.022	0.187	-0.073	0.163	0.015	0.186
industry	-0.045	0.056*	-0.034	0.163	0.067	0.074*	0.041	0.175
working years	0.069	0.187	0.072	0.158	-0.184	0.068*	0.063	0.229
health status	-0.193	0.001***	-0.065	0.084*	0.056	0.036**	0.024	0.283
financial planning	-0.035	0.162	0.092	0.173	0.086	0.184	-0.023	0.087***
risk preference	-0.188	0.088*	0.046	0.128	-0.192	0.155	0.007	0.023**

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 3-16, we can see that there are obvious differences in several financing channels. Generally speaking, wealth and health are important factors affecting household financing choices. The wealthier the household is, the more likely they will borrow money from a bank; the less healthy the household is, the more likely they will borrow money from a non-bank.

First of all, we can see that the factors affecting private borrowing at a significance level of 10% are total assets, children, elderly people, region, education, industry, health status and risk preference. But at the significance level of 1%, we can find that the main factor is health status; meanwhile, it has a negative relationship with private borrowing, and thus it can be said that if a household is in a bad condition of health, it is more likely to borrow money from relatives and friends. The second important factors are region and total assets. Region has a positive relationship with private borrowing. This means that households in the third-tier cities are more willing to borrow from relatives and friends than ones in the second- and first-tier cities. Total assets have a positive relationship with private borrowing. Other factors such as children and elderly people have relationships with private borrowing, but we think the results are not robust enough to comment upon them. Moreover, education and industry factors have different relationships with private borrowing; it can be inferred that households with members who work in the public sector are less willing to borrow money from relatives and friends. The cause may be that the ones who work in the public sector can borrow money from banks more easily than others.

Second, from the comparison of bank and non-bank borrowing, we can find that family wealth is an important factor affecting the two. It can be inferred that at present, to a certain extent, bank borrowing has relatively more financing constraints, which restricts the choices of families. At the same time, we can find that health is negatively correlated with bank borrowing; that is, the better the family health is, the better the choice of bank borrowing is. Health is also negatively correlated with non-bank borrowing; that is, the worse the health is, the more non-bank borrowing is chosen.

Furthermore, region and non-bank borrowing are negatively correlated—i.e., there are more non-bank loans in first-tier cities. This may be related to the convenience of borrowing. At present, most non-bank financial institutions have not yet operated nationwide, and this restricts household choice. And there is a negative correlation between working years and non-bank borrowing. The longer the working years are, the less non-bank borrowing there is. This is consistent with the impact of wealth to a certain extent. Generally speaking, the longer the working years are, the richer the household will be.

Finally, we can see that wealth and health status are the key factors affecting household lending decision-making in China. By analyzing the factors affecting financing choice, it can be seen that Chinese households tend to borrow from relatives and friends, especially in non-first-tier cities and poor health conditions. Meanwhile, there exist obvious financing constraints with bank loans, so this may restrict the choices of households. But generally speaking, this is consistent with the financing decision-making analyzed above. It can be concluded that the preferred financing choice for most households in China is to borrow from relatives and friends, and then from banks.

4. Credit Cards

In recent years, credit cards as payment instruments and financing channels have become popular in our daily life. From 2003 to 2012, the cumulative credit card issuance in China reached 331 million and it is still growing rapidly (up 16% in 2012), with a total credit of 3.49 trillion RMB. At present, per-person possession of credit cards in China has reached 0.16.[1] With the gradual popularization of credit cards, credit cards play an increasingly important role in our daily life and the country's economy.

Although credit cards in China have developed in great leaps and bounds as a special financial product, they are still limited in comparison with the mature credit card market of developed countries, whether in terms of usage ratio or financing rate. Meanwhile, in mature financial markets, consumers conduct a large amount of credit behavior and payment through credit cards. There is still a huge space left for market growth in China, so it can be inferred that the credit card market in the future will grow rapidly for a few years.

At present in China, compared with the rapid increase in the amount of card issuance, credit card holders have low enthusiasm and activeness regarding card use, and there are a large number of "dormant cards" and "zombie cards".[2] The proportion of China's credit card consumption is far below the international level, and the use of financing is even rarer. These phenomena indicate that the consumers have not formed the habit of using credit cards.

Thus, this chapter explores characteristics of the credit card market in China, and it studies different behaviors of consumers in using credit cards as a financing channel and a payment instrument. Through the analysis of card behavior, we intend to provide theoretical suggestions for expanding the credit card market.

---

[1] The information comes from the People's Bank website.

[2] "Dormant cards" commonly refers to inactivated credit cards, and "zombie cards" refers to cards that use only the minimum number of swipes. According to the statistics of the People's Bank of China, about 30% of China's credit cards are dormant cards.

4.1 Descriptive analysis of credit card behavior

This study uses the data from the China Survey of Consumer Finance conducted by the Chinese Financial Center of Tsinghua University in 2011. The sample covers 25 cities and the total number of samples is 5,911. Combined with the variables selected by the research, the number of valid samples is 4,711 after removing missing values, outliers and invalid values. We use section 9—credit card—in our analysis.

Table 4-1 shows the percentages of credit card holdings. As shown in Table 4-1, the proportion of successful applications for credit cards increases with total household wealth. However, from the overall level, the ratio of urban households applying for credit cards is not high, and more than 50% of consumers have not applied for credit cards.

Table 4-1 The percentages of households holding credit cards

Assets (thousand yuan)	0–50	50–100	100–200	200–500	500–1,000	1,000–2,000	above 2,000
Applied and succeeded	16.8	11.75	17.23	16.28	33.82	46.75	62.05
Applied but did not succeed	12.26	8.1	7.67	12.08	12.83	4.83	2.48
Did not apply	70.94	80.14	75.11	71.64	53.34	48.42	35.47

Based on Table 4-1, we explored further the reasons for not applying for a credit card. Table 4-2 shows the reasons for not applying for a credit card. As can be seen from Table 4-2, most households do not apply for a credit card because it is unnecessary and the ratio reaches 34.4%. In addition, 1.6%, 6.6% and 4.7% of the respondents, respectively, do not apply for a credit card due to “high thresholds, complicated procedures”, “could lead to more consumption exceeding the economic capacity” and “security concerns”. The total rate of not applying for a credit card reaches 47.3%.

Table 4-2 The statistics of the reasons for not applying for a credit card

	Frequency	Percentage （%）
Unnecessary	1,621	34.4
High threshold, complicated procedures	75	1.6
More consumption than economic capacity	311	6.6
Security concerns	221	4.7
Sum	2,228	47.3

Among the whole sample, the average number of times using credit cards is 60.76 times per year and the standard deviation is 429.84 times. Table 4-3 shows the frequency of credit card usage. As can be seen from Table 4-3, most credit card holders use them 7–10 times per month, and that ratio is up to10.1%. The second-highest ratio is that of respondents who use their credit cards 4–6 times per month, and that ratio is 9.6%. Some respondents use their credit cards 3 times or less per month. This is sometimes called having a “dormant card”, and it accounts for 9.1% of users. The ratio of credit card holders using their cards 11–15 times per month is only 0.5%, which is the lowest within the five groups. What is interesting is that the number of active credit card users who use credit cards more than 15 times per month is relatively high, and that ratio is 7.1%; and ones who use credit cards more than 15 times per month can be seen as loyal customers.

Table 4-3 The statistics of credit card usage frequency

	Frequency	Percentage （%）
1–3 times	429	9.1
4–6 times	452	9.6
7–10 times	476	10.1
11–15 times	24	0.5
15 times and above	349	7.4

The average monthly credit card usage of Chinese households is shown in Table 4-4. It can be seen that the most monthly usage of credit cards is for less than 3,000 RMB, and that proportion is about 34.2%. The second-most monthly usage is for 3,000–6,000 RMB, and that proportion is about 5.4%. Meanwhile, the percentage of consumers who use credit cards for more than 12,000 RMB per month is about 2.6%, which may be the most valuable consumers.

Table 4-4 The statistics of credit card monthly repayment

	Frequency	Percentage（%）
0–3,000 yuan	1,611	34.2
3,001–6,000 yuan	254	5.4
6,00–19,000 yuan	75	1.6
9,001–12,000 yuan	75	1.6
Above 12,000 yuan	122	2.6

Most households do not use the credit card installment payment function—the proportion was 62.58%. Only 2.83% of consumers often use the installment function of credit cards, which means using the installment function more than five times a year, and 57.31% of the households apply for credit cards only because of convenient payments. Meanwhile, 38.71% of the households use credit cards to get short-term funds for shortages in addition to using credit cards for consumption. Additionally, 3.98% of the households use credit cards only for getting short-term funds for shortages.

Figure 4-1 describes the evaluation of credit card quotas by Chinese households. As can be seen, most households think their credit card quotas are appropriate. For the over-low group, it can be found that most households are concentrated with total assets below 200,000 RMB. Meanwhile, we have explored the repayment methods for credit cards; most households tend to make automatic transfer repayments, and the tendency increases with household wealth.

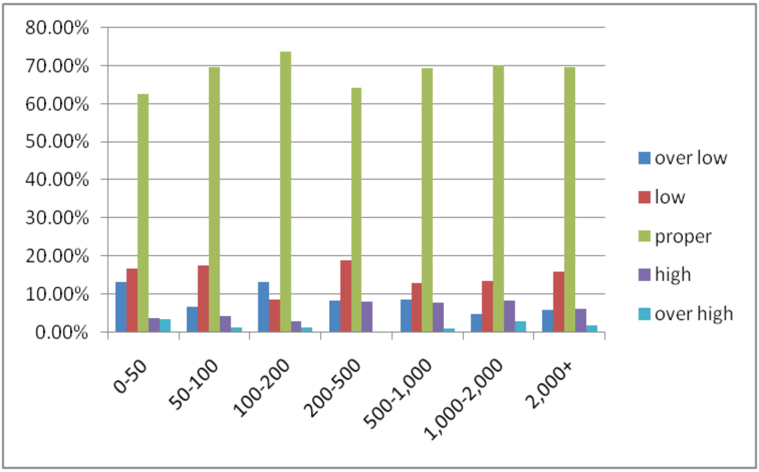

Fig 4-1 Wealth and credit quotas

Figure 4-2 summarizes the situation of deferred credit card payments in China. Here, we define "seldom" as no more than 3 times per year, "sometimes" as 3–6 times per year, and "Often" as more than 6 times per year. Most consumers who hold credit cards have no experience of deferred payment, and that ratio is more than 68%.

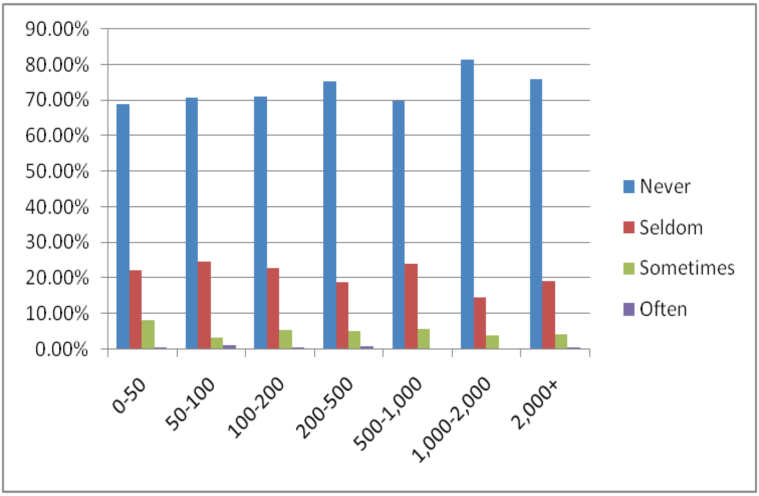

Fig 4-2 Wealth and deferred credit card payment

Figure 4-3 explores the attitudes towards credit cards and consumption. As can be seen, most households believe that credit cards have no stimulating effect on their consumption; even if they have relatively high status of wealth, they also think that there is only a slight stimulation on consumption.

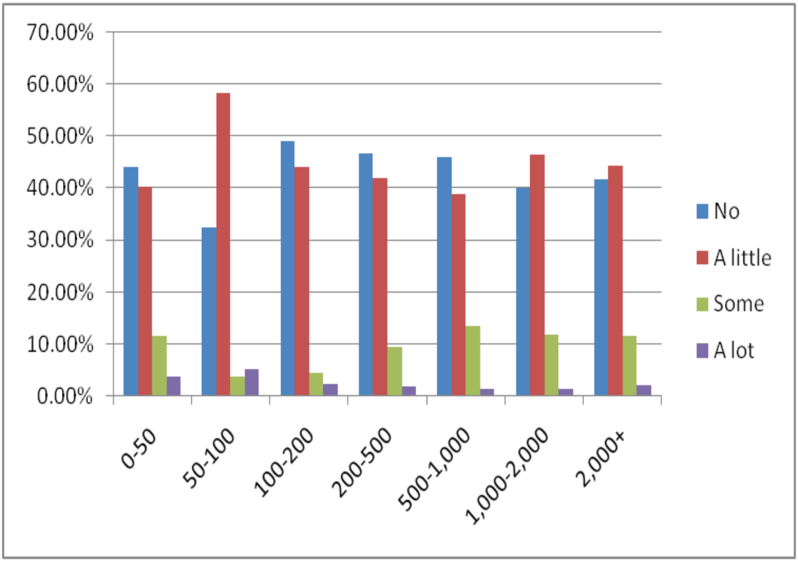

Fig 4-3 Attitude towards credit card effect on consumption

There is relatively high interest with credit card usage. Figure 4-4 explores the knowledge of households regarding credit card interest. From Figure 4-4, we can see that most households know the interest rates of credit cards, and the percentage of respondents whose assets are at the extremes of the wealth levels (less than 100,000 RMB and more than 1 million RMB) know more about this interest.

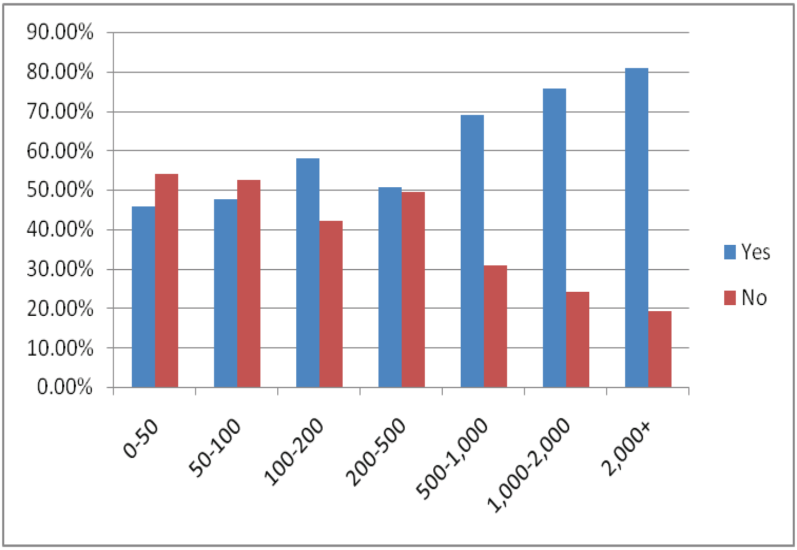

Fig 4-4 Knowledge of credit card interest

Figure 4-5 describes the status of residents' understanding of credit card default. The results show that with the increase of total wealth, the proportion of households who understand the consequences of personal credit card default increases. On the whole, most households are aware of the consequences of personal credit card default.

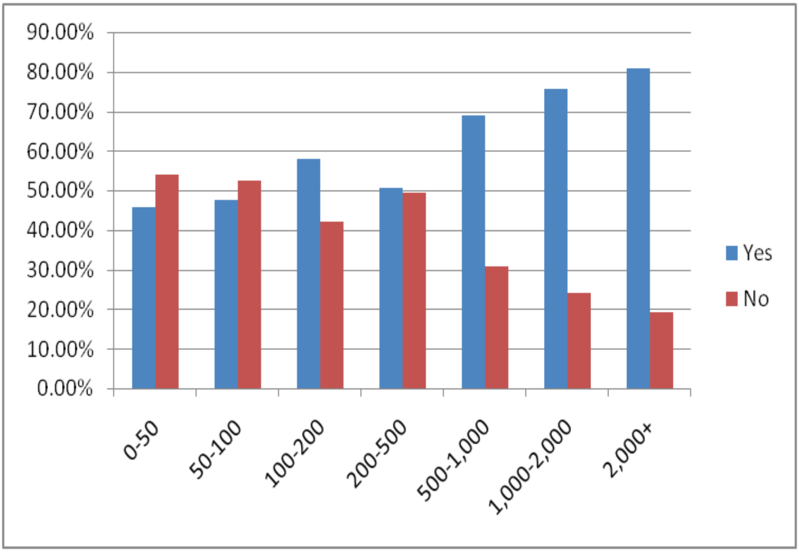

Fig 4-5 Credit card debt default awareness

4.2 Credit card as payment instrument

Credit cards have the dual functions of payment instrument and financing channel. With the increasing of credit card usage in China, payment instruments have played an important role in economics. According to the statistics of the People's Bank of China, the proportion of credit card consumption in total social consumption has increased for 12 consecutive years, and the role of credit cards in stimulating consumption still has space to emerge. Meanwhile, China's banking industry has entered the era of retail; new competition among major banks will be in the field of retail. Thus, as the window to retail financing, the credit card is not only the main tool of consumer credit but it also contains broad prospects for profit space, which has become the focus of commercial banks' competition. A similar industry situation can be seen in the work of Fan, Ji, Wei, and Lambert (2018).

Zhao, Zhao, and Song (2009) claim that card issuers should ignore some missing payments to give more flexible on consumer bill payment, which can differentiate between low-risk, delinquent customers and high-risk customers and gain more benefit. Huh, Chang, Lee, and Lee (2010) explore future cash flows of credit card accounts of Samsung, and they also argue that to tolerate default to a degree will bring more profit for credit card companies. Srikanthverma and Ranaprathap (2011) analyzed the credit card industry in India; they found that the industry has been changing consumer spending patterns. And the credit card market in emerging economies has expanded rapidly in recent years. The total expenditure through credit cards went up to Rs 584.46 billion during 2009–2010. Chen and Lin (2011) use the sample from Taibei to investigate cause and effect among relationship norms, relationship benefit, and relationship value in using credit cards. They argue that a credit card company can gain loyal consumers by encouraging solidarity, which affects functional value via confidence benefits. Steffes, Murthi, and Rao (2011) also claim that "for most financial services firms, including credit card firms, a good customer is also a profitable customer. Managers would like to use marketing tactics that attract the most profitable customers while closely monitoring and perhaps limiting expenditures on marketing tactics that tend to attract relatively less profitable customers. Therefore, managers need to understand the relative effectiveness of different modes of new account acquisition and the impact that the various modes of acquisition may have on overall account profitability". In the research of Parahoo (2012), he found that customer involvement had path loadings of 0.32 and 0.26 on quality and value, respectively, while both service quality and value had direct effects on loyalty, with path loadings of 0.30 and 0.51, respectively. Thus, marketers of credit cards should leverage involvement in their customers by employing strategies such as branding, positioning, and attractive and flexible frequent use benefits to get the customer loyalty. Similar research can be found in the work of Chang and Ijose (2016). Gan et al. (2016) explore credit card holder behavior in China. Singh, Rylander, and Mims (2018) put their focus on the college student's behavior regarding credit cards.

Also, some researchers explore the reasons why consumers don't want to use a credit card. For example, Ricaldi, Finke, and Huston (2013) found that, among credit card users who do not use cards for borrowing (convenience users), rewards are a means to negotiate the implicit price of the interchange fee. Any consumer whose time cost is less than the value of rebates should rationally choose a reward card. Awanis and Chi Cui (2014) define weaker recollections of past credit expenses and overvaluation of available funds—a phenomena the authors call the “credit card effect”. And they promote a measure of susceptibility to credit card misuse and indebtedness (SCCMI). Through their research, they argue that the credit card effect will be another reason for rational consumers to use a credit card. In the research of Chahal, Kaur, and Rani (2015), the authors examine the dimensionality of customer experience in the context of the credit card industry. They also put some focus on why some consumers choose not to use a credit card. Similarly, the research of Wong and Lynn (2017) also claim that credit cards have the easy-money effect.

In short, the credit card market is an important tool to stimulate consumption, but there still a lack of research on the market. Nowadays, the competition among credit card companies is becoming increasingly fierce, so how to find high-quality customers and improve the loyalty of consumers is the core concern of the credit card market. In this section, we focus on the characteristics of customers who use credit cards as a payment instrument, which is the first step in credit lending. In order to find the high-quality customers, we used the data from The Agricultural Bank of China; we explored the background of consumers to define the category. Then we went further with a sample of credit card behavior records to define loyal customers through a special consumption dimension curve.

4.2.1 Data and method

The data used in this section are from the credit card database of a domestic commercial bank. Compared with the data obtained by the questionnaire, the data from the bank have the following characteristics: (1) large amount, which may reduce small sample deviation; (2) high reliability and accuracy. These characteristics also provide a good basis for empirical research. The method of sample selection in this paper is as follows: We selected credit card accounts from February 2007 to February 2008, and we added up the number of card use times of these accounts within one year. After cleaning and sorting out the data, we randomly extracted a sample with 30% proportion, the final number of observations is 838,446.

We measured the frequency of card use within one year as “Cardnum”, which is the dependent variable of this research. Table 4-5 shows the frequency distribution of Cardnum, from which we can see that the frequency is generally not high; a large number of cardholders use their cards less than 10 times a year.

Table 4-5 The distribution of credit card use frequency

Times	Frequency	Percentage（%）	Cumulative percentage
0	55,141	6.58	6.58
1	51,030	6.09	12.66
2	37,411	4.46	17.12
3	36,802	4.39	21.51
4	35,635	4.25	25.76
5	47,533	5.67	31.43
6	43,325	5.17	36.60
7	37,668	4.49	41.09
8	33,731	4.02	45.12
9	30,738	3.67	48.78
10	28,475	3.40	52.18
10+	400,957	47.82	100

On the basis of the frequency of card use per year, we constructed two indicators to measure cardholders' enthusiasm and activity in using credit cards: "Cardact" and "Cardfreq". Cardact is the measure of the cardholder's overall swiping times in one year, and if the user swipes less than 5 times, Cardact will be 0; otherwise, it will be 1 since, in China, if a card user swipes their card less than 5 times, the credit card company will charge a maintenance fee. Cardfreq is an active credit card holder's actual number of swipes in a year, which is the reflection of the cardholder's activity.

Table 4-6 shows the independent variables selected in this research. This research uses three indicators to measure the life cycle factors of cardholders: gender; "marr", which is the marital status; and age. In response to consumer attitudes, we selected two observable indicators to approximate, which are: "Lmt" and "Goldcard". "Lmt" refers to the cardholder's choice of credit limit; 0 means that the cardholder's application quota is higher than the approved quota, and 1 means that the cardholder's application quota is less than or equal to the approved quota. We believe that this indicator reflects the cardholder's eagerness for credit, so it can be used to approximate the creditability of cardholders. "Goldcard" indicates the cardholder's choice of card between gold card and general card; 0 means only gold card, and 1 means that if the gold card is not issued the cardholder can accept a general card. “Goldcard” is also an indicator of credit card holders’ attitudes.

Table 4-6 The summary of independent variables

Variable	Notes	Percentage（%）
Gender	0: male	59.25
	1: female	40.75
Marr	0: unmarried	15.63
	1: married and have child	77.01
	2: married but no child	6.80
	3: other	0.55
Age	0: below 2	0.91
	1: 21–35	49.43
	2: 36–45	36.98
	3: 46–60	12.21
	4: above 60	0.47
Lmt	0: approved > application	49.53
	1: approved <= application	50.47
Goldcard	0: only gold card	48.69
	1: both	51.31

The model of Cardact is shown as equation (4.1), with Cardact as a binary variable, so we apply logistic regression of this model.

$$Cardact = \beta_0 + \beta_1 Lmt + \beta_2 Goldcard + \beta_3 Edu + \beta_4 Income + \beta_5 Gender + \beta_6 Marr + \beta_7 Age + \beta_8 Region + \varepsilon \quad (4.1)$$

The model of Cardfreq is shown as equation (4.2), with Cardfreq defined by Cardnum, so we apply Tobit regression of this model. When Cardnum is more than 5, Cardfreq equals Cardnum; otherwise, Cardfreq equals 5.

$$Cardfreq = \beta_0 + \beta_1 Lmt + \beta_2 Goldcard + \beta_3 Edu + \beta_4 Income + \beta_5 Gender + \beta_6 Marr + \beta_7 Age + \beta_8 Region + \varepsilon \quad (4.2)$$

4.2.2 Empirical Results

Table 4-7 shows the results of credit card use. From Table 4-7, credit card use can be seen from the perspective of the cardholder's life cycle characteristics; traditional life cycle factors (gender and age) are not significant. The possible reason here may be that social and economic factors such as income and occupation have dominant positions. However, another characteristic of the life cycle—the marital situation—has a slightly significant impact on cardholders' card use behavior. The results of the Logit regression show that, compared with cardholders of other marital statuses, the proportion of active cardholders who are married and have children is significantly lower. This is consistent with the traditional life cycle theory. For cardholders who have children, supporting their family leads to a significantly increased burden, and at the same time, they always need to reserve more funds for their children's future, even though income may increase. But people who get married and have children will try to reduce unnecessary expenses, so the willingness to consume is subject to certain restrictions, which shows that the number of children has an important impact on the willingness to consume among Chinese cardholders. Although the influence of the Tobit regression is not significant, the possible reason for the impact of marriage may be that the Tobit regression reflects the cardholder's specific number of uses, so most frequent card users are all unmarried cardholders. In general, the cardholders' behavior in China still has a relatively obvious "life cycle effect".

Table 4-7 The results of using credit cards as a payment instrument

Variables	Logit Coef.	Tobit Coef.
Female	-0.00288	-0.0003
Unmarried	-0.00057	-0.0141
Married with child	-0.0185**	-0.0225
Married without child	0.01	-0.0111
below 20	-0.0008	-0.0011
21–35	0.00301	0.0014
36–45	0.00435	-0.0024
46–60	0.00847	-0.0034

Lmt	-0.00775 ***	0.0003
Goldcard	0.0487***	0.0334***
Graduate and above	-0.0141	0.0091
Undergraduate	0.00167	0.0012
High school and below	0.00564	0.003
50,000 yuan and below	-0.0344**	-0.0780*
50,000–100,000 yuan	-0.0218	-0.0763*
100,000–200,000 yuan	-0.0169	-0.0736*
200,000 yuan and above	-0.0128	-0.0739*
City	0.00957***	0.0113***
Constant	0.8097***	2.2548***

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%,*** indicates that the p-value is less than 1%.

Furthermore, from Table 4-7, from the results of the Logit regression, we can see that the credit limit “Lmt” has a negative impact on cardholders’ card enthusiasm. That is to say, cardholders whose applied-for quotas are higher than the approved quotas are more likely to become passive card users, intuitively. In other words, the cardholders who are more eager for credit are less motivated. The possible reasons for this are as follows: Firstly, cardholders with high applied-for quotas do not necessarily have a positive credit attitude. They are only blindly optimistic about applying for higher quotas, and they do not really have a higher demand for credit funds. This point also reflects the fact that Chinese consumers lack sufficient awareness of credit cards and have greater randomness in credit card applications. Secondly, the traditional research’s explanation for the positive impact of credit attitude on credit card usage is mainly based on the credit card’s lending function; that is, people with positive credit attitudes will use credit cards more, because they are in high demand for credit funds, and credit cards are the most widely used tool for short-term financing. However, in China, credit cards are largely used as a payment instrument rather than a financing channel, so the positive attitude towards credit does not explain the enthusiasm of credit card use. These may be the two main points which distinguish credit cards in China. Moreover, the results of the Tobit regression show that the “Lmt” coefficient is not significant, which further demonstrates that the credit attitude is not robust enough to explain the cardholder’s enthusiasm and activity.

"Gold card" is an important explanation for the cardholder's frequency of use. In the Logit and Tobit regression results, the coefficients of "gold card" are significantly positive; that is, the cardholder who selects "Only gold" beats the one who selects "Both are ok". Both enthusiasm and actual frequency have increased significantly with "Only gold". Although the credit quota of the gold card is higher than that of the general card, considering the coefficient of "Lmt", which has already controlled one's craving for credit funds, we believe that credit attitude is not a reasonable explanation for this result. This phenomenon mainly reflects another remarkable feature of Chinese credit card holders, which is a value-expressive function. It can be thought that if a person has a gold card, it means they have a higher social status. We believe this may be the main cause for the credit card holders' behavior.

Furthermore, the cardholder's income level has an important impact on the frequency of use. In particular, there is a significant negative correlation between the population with the lowest income level and the use frequency. The use frequency of credit cards increases as the income level increases. Generally speaking, respondents with the highest income have the highest card use frequency. The coefficient of the control variable—city, which reflects the urban-rural gap—is significantly positive in both Logit and Tobit regressions. So, cardholders living in cities are significantly more likely to swipe their credit cards than rural cardholders.

Finally, we examine the overall effects of each factor through the Type III Analysis of Effects in Logit and Tobit Regression. Table 4-8 shows the impact of various factors as a whole on the frequency of credit card usage. From the Logit regression results, the cardholder's choice of gold card has the most significant impact on the cardholder's overall card activity, followed by the cardholder's marriage and child statuses. The influence of the city and the income of the cardholders are also obvious. The results of the Tobit regression are generally consistent with the Logit, with a slight difference in that the "Lmt" has no impact on the users' behavior.

Table 4-8 Type III Analysis of Effects in Logit and Tobit Regression

variables	df	Logit		Tobit	
		Wald Chi-Square	P	Wald Chi-Square	P
Gender	1	1.2061	0.2721	0.0076	0.9307
Marr	3	12.8265***	0.005	7.3257*	0.0622
Age	4	0.7556	0.9443	1.7198	0.7871
Lmt	1	8.6898***	0.0032	0.0105	0.9186
Goldcard	1	342.8434***	<.0001	108.5560***	<.0001
Edu	3	1.8942	0.5946	1.1336	0.769
Income	4	9.9557**	0.0412	4.1802	0.3822
City	1	11.8966***	0.0006	11.0495***	0.0009

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

4.2.3 Conclusion

Wilcox, Block, and Eisenstein (2011) state that outstanding credit card debt increases spending for consumers with high self-control. This effect can be eliminated by increasing the available credit on the credit card. Through the analysis of the factors affecting the enthusiasm and activity of credit card holders in China, the following empirical conclusions are obtained: Firstly, from the perspective of life cycle characteristics, age has no significant influence on the cardholder behavior of Chinese cardholders. However, the impact of the marriage situation and whether the cardholder has children is very significant. The cardholders who are married with children are significantly less active than other cardholders. Second, consumer attitudes have a significant impact on the frequency of credit card use in China. The positive attitude regarding status greatly promotes the cardholder's card activity, but the traditional credit attitude which is defined by quota cannot reasonably explain the cardholder's card behavior in China; meanwhile, the cardholder's value expression is the most important key indicator. In addition, the cardholder's usage frequency increases with the increase of income, and urban-rural differences also have a significant impact on the cardholder's usage frequency.

So, for commercial institutions and companies, it is necessary to adopt more effective market segmentation and marketing strategies to expand the market. On the one hand, to encourage customers to form active card habits, commercial companies need to emphasize the convenience and value-expressive characteristics of credit cards. On the other hand, commercial companies should specifically be aware of consumers' marriage statuses and incomes, especially to make reasonable judgments regarding the cardholder's life cycle stage, and then correspondingly develop products and expand the market. And in order to stimulate consumption, the government needs to increase the availability of credit card knowledge and improve the construction of credit card infrastructure, especially to promote credit card acceptance in rural areas.

4.3 Credit card as financing channel

A credit card can be used as a financing tool for most families. With the help of a commercial bank in China, Wang, Lu, and Malhotra (2011) studied consumer credit card debt behavior in correlation with demographics, attitude, personality and credit card features factors. According to the research of Druedahl and Jørgensen (2018), there is a puzzle in credit card debt. Current debt may soften household borrowing constraints in the future, thus providing additional liquidity, but for some median net liquidity, they want to use credit card debt, even if the interest rate is much higher than the return on assets. Similarly, Choi and Laschever (2018) use data from the Health and Retirement Study and examine the role of noncognitive skills as well as the economic, financial and demographic factors that affect the credit card debt. Keys and Wang (2019) claim that 29% of accounts in the US make credit card payments at or near the minimum payment. And they find that anchoring to a salient contractual term has a significant impact on household credit card repayment decisions. Also, there are some studies that put their focus on college students' behavior with credit card debt, such of those of Norvilitis et al. (2006), Joireman, Kees, and Sprott (2010), and Paul, Nolan, and Smith-Hunter (2017).

In this study, we put our focus on the credit card financing channel, and we try to find the answer to the question: "Who will use the credit card debt?" The data used in this section are from the credit card database of a domestic commercial bank, as above.

4.3.1 Variables

The dependent variable is divided into two stages. The first stage examines whether the cardholder will become a revolver; that is, the possibility of the cardholder using a credit card to borrow. The dependent variable used here is the binary variable called "Carduse". When the cardholder is observed to borrow or generate interest in the account, the value refers to 1; otherwise, if the cardholder account is repaid totally in each period or only uses the credit card as a payment instrument, the value is 0. The second stage examines the degree of cardholder overdraft, and a numeric variable known as "Carddebt" is used here. "Carddebt" represents the total amount of overdraft interest generated by the cardholder during each period, which can reflect the total amount of interest within a certain period of time in order to measure the total overdraft of cardholders.

We use age and marital status to measure the cardholder's life cycle. Age is expressed in terms of the natural logarithm of the cardholder's age, with the aim to capture the non-linear relationship between age and credit card overdraft. Marital status (Marr) is a nominal variable, which is separated into "unmarried", "married", "without child" and "with child".

The socioeconomic status variables include: income is the natural logarithm of annual income. Job lever (Job) is defined as junior, intermediate and senior positions, and they are represented in turn by the numbers 1, 2 and 3.

Other credit variables are credit limit (Lmt), percentage of total amount (Ultrate) and account age (AccAge). Lmt is the maximum credit amount that the cardholder can obtain. From the perspective of supply and demand of lending behavior, Lmt reflects the supply of credit. Therefore, it is the upper limit of credit card debt of the cardholder. Ultrate is the ratio of the average amount of card usage divided by the total credit amount during the observation period (including consumption and cash withdrawal). AccAge is the number of months since one opened one's credit card. The summary of discrete independent variables can be seen in Table 4-9.

Table 4-9 The summary of independent variables

Variable	Notes	Percentage (%)
Gender	0: male	59.25
	1: female	40.75
Marr	0: unmarried	15.63
	1: married and have child	77.01
	2: married but no child	6.80
	3: other	0.55
Job	1: Junior	15.12
	2: Medium	48.31
	3: Senior	36.57
Lmt	0: approved > application	49.53
	1: approved <= application	50.47
Ultrate	1: under 10%	17.18
	2: 10%–20%	22.54
	3: 20%–30%	42.69
	4: above 30%	17.59

4.3.2 Model and results

We used the Logit regression model to analyze the factors which influence credit card debt, as in equation (4.3).

$$\begin{aligned} log\, Carduse = \beta_0 + \beta_1 \ln Age + \beta_2 Gender + \beta_3 Marr + \beta_4 Job \\ + \beta_5 Lmt + \beta_6 Ultrate + \beta_7 Accage + \beta_8 Income + \varepsilon \end{aligned} \tag{4.3}$$

And then, we explored the total amount of credit card debt with linear regression, as in equation (4.4).

$$\begin{aligned} Carddebt = \beta_0 + \beta_1 \ln Age + \beta_2 Gender + \beta_3 Marr + \beta_4 Job \\ + \beta_5 Lmt + \beta_6 Ultrate + \beta_7 Accage + \beta_8 Income + \varepsilon \end{aligned} \tag{4.4}$$

The results of these two models can be seen in Table 4-10.

Table 4-10 The results of credit card as financing channel

Variables	Logit Coef.	Linear Coef.
Female	-0.0134***	-1.2702**
Unmarried	0.1218*	0.2316**
Married with child	-0.1165**	-0.2334*
Married without child	0.0431	-0.1832*
lnAge	-0.3528***	-0.3894**
Job_1	0.0745*	0.1123
Job_2	0.0412	-0.0072
Income	-0.6237	-0.5634*
Lmt_1	-0.0019*	0.0164
Ultrate_2	0.0182	0.2314*
Ultrate_3	0.0253	0.1863
Ultrate_4	0.1863*	0.3764*
AccAge	0.1975***	0.2431**
Constant	0.8097*	0.7254**

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%,*** indicates that the p-value is less than 1%.

We can summarize the results in Table 4-10 as the following points. First, we can see with the life cycle that females are less likely to have credit card debt, compared with male ones. Second, the age of the cardholder has a significant impact on the likelihood of credit card revolving liabilities and the degree of debt. The younger the cardholder is, the more likely they will hold credit card liabilities; and the younger the cardholder is, the higher the amount of credit card debt is. So there is a negative relationship between age and credit card debt. Third, the marital status also has a significant impact on credit card revolving liabilities. Unmarried families are most likely to have credit card debt, and married families with children have the least credit card debt. This is a little different from the traditional life cycle theory, which supposes that families with children will spend more money and might use credit card debt more than others. We believe that, in China, most households have the awareness of saving for the next generation, so this will reduce the possibility of using credit cards as a financing channel. To sum up, China's cardholders' credit card financing behavior shows obvious life cycle effects. Cardholders of different ages and marital statuses show significant differences in credit card liability behavior and debt level.

Moreover, job does not have a significant impact on credit card use as a financing tool; income only affects the credit card's debt level and it presents a negative correlation—ie, as the income increases, the credit card debt will decrease.

In terms of credit card characteristics, we can find that both the amount of credit card applications and the amount of final approvals have no effect on whether household uses credit cards as financing channel or not. However, it can be found that as credit card usage increases (Ultrate), families have more and more credit card liabilities. Meanwhile, the age of credit card accounts is positively related to credit card liabilities. This is easy to understand. When a household has a better understanding of the function of credit cards, this will increase the possibility of using a credit card as a financing channel.

4.3.3 Discussion

We went further with the credit card debt. As we described above, the credit card as a financing channel was measured as a binary variable for using it as a financing channel or not. . As can be seen, some credit card debts are caused by negligence, such as forgetting the repayment date. And these cannot be on behalf of the usage as the financing channels for credit cards. Therefore, we further discuss the financing channel of credit cards to limit the dependent variables to installment payments. If one has used the installment payment, then the variable installment will be marked as 1; otherwise, 0. We use the same independent variables for this hypothesis, and the result can be found in Table 4-11.

Table 4-11 The factors influencing installment repayment

Variables	Logit Coef.
Female	-0.0283*
Unmarried	0.2521**
Married with child	0.1124
Married without child	0.1032
lnAge	-0.1673**
Job_1	0.0672
Job_2	-0.0432

Income	-0.3865**
Lmt_0	-0.0004
Ultrate_2	0.1123
Ultrate_3	0.1653**
Ultrate_4	0.3525***
AccAge	0.3421***
Constant	-0.5097*

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%,*** indicates that the p-value is less than 1%.

It can be seen that most of the conclusions are consistent with credit card liabilities as above. Besides, it is worth noting that Job and Lmt are no longer significant indicators with installment repayment. The coefficient of "Ultrate" is significantly increased. It can be inferred that a household is more likely to have an installment repayment with the increasing of credit card use. So, generally speaking, if a company wants to expand the credit card installment market, the best way is to increase usage and try to keep the consumer.

4.3.4 Conclusion

Based on the micro-data from the credit card centers in commercial banks, this research empirically analyzed and discussed the factors affecting credit card revolving liability behavior from the aspects of life cycle, socioeconomic status and credit card characteristics.

The main conclusions are as follows: First, China's cardholders' credit card revolving liability behavior has a significant life cycle characteristic. Female cardholders are more conservative regarding credit card liability than male cardholders. Young people are more likely to hold credit card revolving liabilities than older people, and the amount of debt is larger than that of older people. At the same time, compared with unmarried cardholders, cardholders who are married and have children are significantly more conservative in their credit card liability behavior. Second, socioeconomic status has a slightly significant impact on the likelihood of credit card liabilities and the level of debt. With increasing income, the amount of credit card revolving liability will decrease. But there is no significant relationship between job level and revolving liability behavior. Third, the credit card limit has no obvious effect on credit card liability, but the age of the credit card account is one of the most important factors affecting the cardholder's cyclical debt behavior. The possibility of cardholders using revolving liability increases with the account age. Meanwhile, the total amount of credit card usage is another important factor affecting revolving liability; household revolving liability increases as credit card usage increases.

With the gradual popularization of credit cards in China, the role of credit cards in China's economic development has become increasingly important. In order to make credit cards as a financing channel more acceptable, more requirements need to be fulfilled: ie, one must establish a whole market with all kinds of consumer credit products covering the whole life cycle, so that the household can get more distinct information. And, in China, a commercial bank cannot approve credit cards to people over the age of 60; this leaves a large market for these elderly people. The other recommendation is to improve the credit card services. According to the survey in the first section—the reasons for customers not to choose credit cards—about 1.6% of households think the service is not convincing enough.

Combined with the empirical conclusions, we can get the following insights: First, the management of credit cards in commercial banks should change to pay more attention to consumer attitude. Commercial banks need to realize that the credit demand of households varies with different life cycle stages, so there is a big space to design the credit products based on consumer needs. Second, it is always a good to improve the enthusiasm of cardholders, which can increase not only the usage of payment but also the amount of revolving liabilities. Furthermore, the commercial banks should improve their services; as can be seen, with the increase in account ages, more and more cardholders will have revolving liabilities. So, it can be said that the longer the cardholder uses the credit card, the easier it is for them to use the revolving liability. Therefore, commercial banks need to strengthen customer relationships and establish long-term relationships with customers. This will not only help to cultivate cardholders' good credit card usage habits, which may help to improve the bank's profitability, but also facilitate the management of credit card default risks.

5. INTERNET FINANCING

Research on the traditional credit market concentrates on business financing behavior. However, due to the advance of modern internet technology, more consumers choose borrowing from the internet. Bypassing banks, internet financing is a special type of credit market in which individual borrowers take out microloans online without collateral or intermediation from financial institutions. There are two main types of internet financing in China. One is P2P (person to person), in which individuals lend money to individual borrowers directly; the other is P2C (product to consumer), in which individuals invest on products, and the products are the loans from borrowers.

In China, P2P platforms have gained popularity and market recognition in recent years. The first internet lending platform was established in 2007. The internet financing industry explosion began in 2013, when there was a surge in the number of online platforms. Approximately 150 platforms were set up in 2013, accounting for 50% of the total number of internet financing platforms in China. The internet financing sector continued to mushroom in2014; about 900 platforms were set up in 2014. More than 2,000 platforms were in operation by mid-2016, with outstanding loans reaching 209 billion RMB. Credit Ease, a major P2P lending company, was established in 2010, enabling individuals with surplus funds to lend to others who want money. Like the Lending Club in the US, Credit Ease is the biggest internet financing platform in China.[3]

In 2013, more than 200,000 people lent a total of 105.8 billion RMB on approximately 800 internet financing lending platforms. In 2014, there were 1.16 million investors and 630,000 borrowers involved in the internet financing sector. Compared to 2013, this means 364% and 320% increases, respectively, for participating investors and borrowers on a yearly basis. More than 1.5 million internet financing platforms in China are currently involved in matching lenders and borrowers. Loans via internet financing platforms reached 149.55 billion RMB in 2017, which was1.86 times the transactional volume in

[3] http://www.01caijing.com/article/34857.htm

2016. But internet financing of all types still only comprises a tiny fraction of the 65 trillion RMB in outstanding loans in the formal banking system (Wang 2017).

Research on internet financing in China is limited but increasing. So far, most studies in the Chinese context focus on the description at a macro level. Although researchers describe characteristics of the credit market, systematic and theoretical explanations are limited. Internet financing, as an emerging consumer financing market in China, is particularly under-studied. Although it has greater convenience, many consumers still prefer to borrow from banks, given the same interest rate. In this chapter, we aim to explore the attitude of households towards internet financing; then, we analyze the behavior of investors who engage in lending through the internet and what products interest the investors; finally, we discuss the borrowers who can get loans from the internet successfully and the factors that influence the deals.

5.1 Attitude towards internet financing

According to the white paper of the China P2P lending service industry, there are several advantages of online borrowing over traditional borrowing. Firstly, the online borrowing process is simpler than traditional borrowing; using the internet for data transmission and customer filtering, consumers can get loans more conveniently. Secondly, through non-manual audits, credit risk is assigned more objectively. Finally, the process of borrowing can break through time and space constraints. Though there are many advantages with P2P borrowing, not every consumer knows about it. In 2013, China's government enacted new a regulation and policy to actively support private capital to enter into the financial sector. Internet companies take an active role in this new field, combining their competence in e-commerce and social network domain. However, there are unpredicted challenges for internet financing as it is still not supervised and regulated by the central government (Jia, Ma, and Katzy 2014).

So there is huge gap between the practice and research. With the gradually increase in the internet financing market, more and more researchers wonder what the true attitude towards internet financing is. Why do people not want to take part in the convenient service? Is there any bias of information? In this part, we focus on the attitude of households towards internet financing; using the survey data, we try to answer the question about the concerns with internet financing.

The most commonly used variables in analyzing financing behavior are demographic variables that are discussed in the life cycle model proposed by Ando and Modigliani (1963). In their research, “Demographic variables generally refer to the basic characteristics of a person, including gender, age, marital status, whether they have children, life cycle stage, etc. The life cycle hypothesis assumes that rational consumers aim to maximize the utility of their whole life. The theory holds that in spite of the constant changes of personal income, families tend to apply financial instruments to achieve a stable consumption flow in the life cycle and the income consumption ratio is unchanged. Therefore, the life cycle hypothesis is often employed as an important basis for classifying household groups.” However, there is still a fierce debate among researchers on what is the precise definition of life cycle stage, except for the essential variables—age, marital status and children.

Social stratum is an overall measure of the social status of individuals or families based on economic conditions and educational background. Cohn and Vaccaro (2006) showed that people with above-average incomes and at least some high school education were more likely to raise capital via new payment methods than those with below-average incomes and less-than-high-school education. Debt condition directly affects families’ credit behaviors. Johnson, Ashta, and Assadi (2010) discovered that there existed a strong connection between the debt and the usage of P2P borrowing. Chen, Huang, and Ye (2018) gave some remarks on individual investment and borrowing decisions considering platform regulatory conditions in China, and argued that the middle-income households who had some level of debt would be more willing to borrow. Song, Chen, Zhou, and Wu (2018) checked the performance of P2P online lending platforms in China.

5.1.1 Data and variables

In our research, we used the survey data to explore the attitude towards internet financing. Our survey questionnaire was extracted from the one used in the China Survey of Consumer Finance conducted by the Center of Tsinghua University in 2011 and the data collected from the student families of a major research university in Beijing as a side survey during the data collection for the third wave of the China Household Finance Survey in 2015.

The original data set included 1,011 household samples. In order to guarantee samples' qualities, we excluded respondents under 18 years old as well as samples with incomplete answers. Finally, 989 valid household samples were used in this study.

We used life cycle variables and social stratum variables as the independent variables. We separated households into two groups: one including those who have successfully applied for internet financing, and the other including those who have not gotten it. The summary of the variables can be seen in Table 5-1.

Table 5-1 The Summary of Independent Variables

Class	Variables	Question	Type
demographic	gender	1A	binary variable
	marriage	1B	discrete variable
	age	1C	continuous variable
	education	1G	discrete variable
	cash	2H	continuous variable
	income	2I	continuous variable
risk attitude	Ara	2C and 2D	continuous variable

The dependent variable is internet borrowing behavior based on the survey question: "Have you ever borrowed money via the internet?" where 1 refers to the ones who borrowed successfully, and 0 refers to the others. In the survey, consumers answered this question using 2 to refer to "apply but not succeed" and 3 to refer to "did not apply". We coded all of these as 0, indicating that they are not used. It is known that the dependent variable is a dummy variable, and the normal error cannot directly explain it. We defined n_i as the number of answers 1, then we used $\theta_i = \frac{n_i}{y_i}$ to fit the logistic regression as $log(\theta_i) = ln(\frac{\theta_i}{1-\theta_i})$. We labeled this variable as "Interfin".

Another dependent variable relates to the most important indicator for those who have not applied for internet financing on the survey. The question is: "What are you most concerned about regarding borrowing via the internet?" where 1 refers to "high interest"; 2 refers to "more collateral"; 3 refers to "time line of repayment"; 4 refers to "personal privacy"; and 5 refers to "other factors". From the data statistics, we can find that 36.45% of the households' answers are 1, 21.17% of them are 2, and the others are 18.33% for 3, 15.24% for 4 and 8.81% for 5. So it can be inferred that most households believe borrowing via the internet will charge them a high interest rate. We label this variable as “finconcern”.

For the ones who have applied for borrowing via the internet and succeeded, we can see their attitude through the question:“What is the most important indicator when you select a financing platform for borrowing?” where 1 refers to “relatively low interest”; 2 refers to “short time to wait for money”; 3 refers to “satisfying the total quota”; 4 refers to “convenient to apply”; 5 refers to “extension available”; 6 refers to “no need for collateral”; and 7 refers to “other factors”. Similarly, the statistics for the seven categories are 9.24%, 26.17%, 28.44%, 18.22%, 6.16%, 7.25% and 4.52%, respectively. We labeled this variable as “Seleindicator”.

5.1.2 Results and discussion

The logistic regression of “Interfin” can be found in Table 5-2. Table 5-3 lists the results for the most important factors. Table 5-4 is the ordered logistic regression result of the most important indicator when one selects the internet financing platform.

Table 5-2 The result of factors that affect internet financing

Variables	Logit Coef.
Female	-0.0516*
Unmarried	0.0724**
Married with child	-0.0124*
Married without child	0.0538*
lnAge	-0.0973*
Edu_high school	0.0773
Edu_college	0.1132
Income	-0.2442***
Cash	-0.3024**
Ara_2	0.1824
Ara_3	0.1913
Ara_4	0.2672**
Ara_5	0.2981***
Constant	-0.4124*

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

As can be seen from Table 5-2, females are less willing to use internet financing, and families with children are less willing to use internet financing. People's willingness to use internet financing declines with increasing age. At the same time, it can be inferred that the main benefit of internet financing is to solve the shortage of funds, so the results from Table 5-2 also verify this point. When the family income and cash are increasing, the willingness of the family to use internet financing decreases obviously. Also, the result is consistent with the risk attitude; the households who have higher risk preference will engage in internet financing more than others.

Table 5-3 The results for households' concerns regarding not applying for internet financing

Variables	Choice 1 Coef.	Choice 2 Coef.	Choice 3 Coef.	Choice 4 Coef.	Choice 5 Coef.
Female	0.1321	-0.1183*	0.0027	0.0178*	0.1533
Unmarried	0.1124*	-0.0254	0.0084*	-0.0244	0.1024*
Married without child	0.1825*	0.0342	-0.0035*	0.0133	0.0795
Married with child	0.2231*	0.0377	0.0012	-0.0169*	-0.0693
lnAge	0.2273**	0.0897	-0.0026	-0.0523**	0.0018
Edu_high school	0.1423	0.0672	0.0053	-0.0157*	0.0116*
Edu_college	0.1462	0.0155	-0.0072	-0.0132	0.0024*
Income	-0.0834***	-0.1834**	-0.0899**	0.0162*	0.0133*
Cash	-0.0699**	0.5352***	0.2437***	0.0148**	0.1821**
Ara_2	0.0652	0.0562**	0.1246**	0.2452**	0.1523
Ara_3	0.0733	0.0388*	0.1157*	0.2279**	0.1655*
Ara_4	0.0541	0.0279*	0.0892*	0.1321	0.1854*
Ara_5	0.0859	0.0134	0.0014	0.1832*	0.1239
Constant	0.1721*	0.1822*	0.1937	0.1896	0.1743*

Note: * indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

It can be seen from Table 5-3 that families have largely different concerns about internet financing. First, married people have more concerns about high interest rates, and seniors are also more concerned about this. On the contrary, households with higher incomes and more cash are less concerned about high interest rates. Second, women's concerns about collateral are an important reason for not applying for internet financing, and there is an interesting finding: the higher the household income is, the less concerned it is about collateral; the more cash a household has, the more collateral will be considered. In terms of risk attitudes, it can also be found that as the degree of risk aversion decreases, household concerns about collateral will increase. Third, for the time line for repayment, unmarried families pay more attention to this aspect, while married families without children pay less attention to this. Families with higher incomes pay less attention to this aspect, and obviously families without cash pay more attention to this. But there is no obvious rule with risk attitude. For the personal privacy concern, we can see that women have more privacy concerns, while married families with children have fewer privacy concerns, and privacy concerns decrease with age. Persons with high-school-or-below education have fewer concerns about privacy. At the same time, there is a positive correlation between income, cash and privacy concerns.

Table 5-4 The results for households' choices regarding internet financing platform selection

Variables	Choice 1 Coef.	Choice 2 Coef.	Choice 3 Coef.	Choice 4 Coef.	Choice 5 Coef.	Choice 6 Coef.	Choice 7 Coef.
Female	0.1252	-0.1432*	0.0087	0.0892	0.0038	-0.0079*	0.0113
Unmarried	0.1132	0.0243*	0.0122*	0.1133	0.1197*	-0.0124*	0.0825
Married without child	0.17425	0.0031	-0.0163	0.1452	0.1203*	-0.0098*	0.0744
Married with child	0.1851*	0.0019	0.0086	0.2688**	0.1283**	0.0154**	0.0126
lnAge	0.1371*	-0.0842*	-0.0058	0.0927*	0.0078	0.0122*	0.1842*
Edu_high school	0.1523	0.1153*	0.0124	0.0342	0.0164*	-0.0122	0.1129*
Edu_college	0.1162	0.0726	0.0092	0.0014	0.0112	-0.0099	0.1092*
Income	-0.1226**	-0.0042*	-0.0065**	0.1211	0.0083	-0.0142*	-0.1035*
Cash	-0.3644***	0.6624***	0.7433***	0.1534*	0.2341**	0.6722***	0.1621**
Ara_2	0.0823**	0.0421*	0.1534**	0.1173**	0.0960*	0.0328*	0.0732
Ara_3	0.0644**	0.0432*	0.1627*	0.1282	0.0922	0.0214*	0.0574*
Ara_4	0.0631	0.0466*	0.0893*	0.1449	0.0914	0.0011	0.0632
Ara_5	0.0811	0.0524*	0.0826	0.1446	0.0933	0.0009	0.0521
Constant	0.1824*	0.1755*	0.1861*	0.1885	0.1643*	0.1621	0.1726

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%,*** indicates that the p-value is less than 1%.

From Table 5-4, we can infer some cues. Firstly, females have a significantly different choice of internet financing platform than males in terms of short time line and collateral. It seems females have negative attitudes towards short time lines and collateral; they may be careless about these factors. Secondly, in different life cycle stages, unmarried households have more concerns about short time lines and convenience and fewer concerns about collateral; married-without-child households have more concerns about convenience and fewer concerns about collateral; and married-with-child households have more concerns about low interest, short time lines, convenience and extension. So, it can be seen that if a household has children, its members will think more about the choice of internet financing platform. The age variable has similar conclusions. Thirdly, educational variables have no significant impact in most cases. Contrary to the intuitive perception, families with less-than-high-school education have more concerns about the choice of internet financing platform. Fourthly, income and cash variables affect most choices. It can be summarized that income is negatively correlated with low interest rates, short time lines, total quota and collateral. But cash is positive with nearly all of these choices, except low interest rates. So, to sum up, it can be said that cash decides the internet financing choice, but income makes a compromise which can partially remove low interest rates, short time lines, total quota and collateral from the consideration. Finally, for risk attitude, the more risk preference there is, the fewer considerations there are. For risk preference, the final factors to be considered are short time line and total quota.

In summary, we can see that life cycle is a key factor which has great impact on internet financing behavior. At the same time, income and cash are important factors affecting the choice of internet financing platform, but they have different impacts on the choice of internet financing platform. Generally speaking, cash makes households consider time and convenience more important; and income makes extension and other factors more important. It is interesting that risk attitudes do not play a key role in the choice of internet financial platforms. We believe that maybe most of the people who are involved in internet financing have risk preference; they may not be sensitive to risk, this is consistent with the results of who engages in internet financing.

5.2 Investors' behavior

In this section, we put our focus on who will lend money on the internet. In the research of Feng, Fan, and Yoon (2015), the authors investigate key factors affecting lenders' bidding strategies using three measurements for the popularity of loans: funding success, number of bids and funding time. Mild, Waitz, and Wöckl (2015) argue that many loans are not secured by collateral in online lending; the assessment of the creditworthiness of the borrower is the most important task. They think that not all investors can transform information to proper market activities, and they claim that the quantifiable banking data is the key to information transform. Meanwhile they put forward a decision support tool to help others to predict default risk. Chen, Jin, Zhang, and Yang (2016) address the research on investor decision-making behaviors in P2P lending from the perspective of rationality and sensibility, not only to more thoroughly examine the factors affecting P2P lending but also to contribute to P2P platform builders' and investees' knowledge. In recent researches, it can be seen that if one has the ability to assess the default risk of different loans, one can make more effective allocation of his money across different P2P markets. So the research of Guo, Zhou, Luo, Liu, and Xiong (2016) puts forward a data-driven investment decision-making framework for this emerging market that is based on an instance-based credit risk assessment model. Caldieraro, Zhang, Cunha, and Shulman (2018) propose and test a theory in which counter signaling provides a mechanism to attenuate information asymmetry about financial products (loans) offered on the internet financing platforms in China. And nowadays, the data types online platforms use are complex and involve unstructured information such as text, which is difficult to quantify and analyze. Loan default prediction faces new challenges in P2P lending. Jiang, Wang, Wang, and Ding (2018) propose a default prediction method for P2P lending combined with soft information related to textual description. From these studies, we can know that the key for investors regarding online investment is to judge the borrowers' credit. So how to judge one's credit is the key to the problem.

Also, there are some studies that focus on the trend of internet financing, which can be found in the work of Stern, Makinen, and Qian (2017), Zhao et al. (2017), Song, Chen, Zhou, and Wu (2018), and Havrylchyk and Verdier (2018). In these studies, the authors all agree that internet financing is playing a more important role in China now.

And how to analyze the behavior is one of the most important questions which can play a key role in the development of this industry.

So, in this part, we explore the behavior of investors and, through the data from Credit Ease, we try to give an answer to the question of who will invest online, and what is the key factor which will make the investors invest.

5.2.1 Data and variables

The data in this research came from Credit Ease in China. There are a total of 178,000 detailed observations of investment from lenders during February 2008 to October 2011. Table 5-5 shows the details of the lenders, and Table 5-7 shows the behavior attributes of the lenders' investment.

Table 5-6 The description of investor variables

Variables	Description	Notes
LENDER_ID*key	Investor Id	
CARD_TYPE	Id Card type	
GENDER		0:male;1:female
TOTAL_AMT	Investment in last three months	1: below 300,000; 2: 300,000–1,000,000; 3: above 1,000,000
PAYMENT	Payment	1: through bank; 2: through Credit Ease; 3: Other
CONTRACT		0: nonfixed investment; 1: fixed investment; 2: both
RECOMMENDED		0: none; 1: employee; 2: consumer; 3: other
CITY		1: tier-1 cities; 2: tier-2 cities; 3 tier 3- cities; 4: other
BIRTHDAY		

Table 5-7 The description of investor behavior variables

Variables	Description	Notes
LENDER_ID*key	Investor Id	
PRINCIPLE_AMT	Investment amount at the beginning of the period	
REGULAR_AMOUNT	Monthly investment	
ASSET_VALUE	Asset value	
PLAN_INVEST_AMT	Investment in plan	
NEW_INVEST		0: No; 1: Yes
STATUS		0: hold on; 1: sell part; 2: sell out; 3: in archiving; 4: in risk; 5: cancel
REGULAR_DATE	Fixed investment effective date	
REGULAR_END_DATE	Fixed investment expiring date	
DUE_REINVEST_AMT	Amount of revolving investment	

5.2.2 Empirical study

In this section, we focus on the investors' behaviors. Since there are obvious differences among investors within different cities, we divided the investors into groups using the city category and then analyzed the behavior of the investors. We calculated "regular" as the period of investment, which is calculated from regular_end_date minus regular_date; and we calculated age using birthday.

It can be seen from Tables 5-6 and 5-7 that there are many variables in investors' behavior, and there is a clear correlation among them. Therefore, we first used the factor analysis method to determine investors' investment behavior factors. The results of the factor analysis can be found in Table 5-8.

Table 5-8 The results of factor analysis for investment behavior

Factors	Variance	Percentage	Cumulative
1	3.611	51.583	51.583
2	1.08	15.431	67.015
3	1.006	14.372	81.387
4	0.903	12.899	94.285
5	0.396	5.659	99.944
6	0.003	0.046	99.99
7	0.001	0.01	100

As can be seen from Table 5-8, three factors account for more than 80% of the explanatory variance. So we take three factors into consideration. The translation of factors can be found in Table 5-9.

Table 5-9 The results of factor translation

	1	2	3
age	0.201	0.467	-0.645
principle_amt	0.952	-0.035	-0.02
regular_amt	0.942	0.004	0.09
asset_value	0.542	-0.031	-0.024
regular	0.049	0.787	0.09
plan_invest_amt	0.681	0.005	0.089
due_reinvest_amt	-0.001	0.491	0.752

From Table 5-9, we can see that the fixed investment amount is the main explanatory component in the first factor; the fixed investment period is the main explanatory component in the second factor, and the reinvestment amount is the main explanatory component in the third factor.

Then, we put cities into groups and explored the factor scores based on city group. The results can be found in the Figures 5-1, 5-2 and 5-3. Please note that we only plotted eight cities' results in tier 2 and 10 cities' results in tier 3, and these cities were selected randomly.

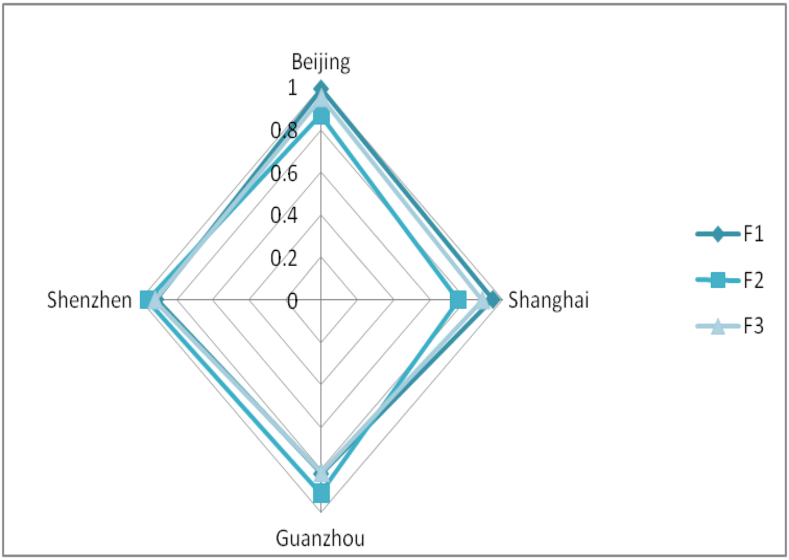

Fig 5-1 The investment scores in the first-tier city group

As can be seen from Figure 5-1, investors in first-tier cities are more enthusiastic about internet investment. Among them, Beijing has obvious advantages in terms of fixed investment amount, and Guangzhou is slightly inadequate; Shenzhen's fixed investment period score is the highest, while Shanghai's score is relatively low; and Beijing scored the highest in terms of reinvestment, while Guangzhou scored the lowest.

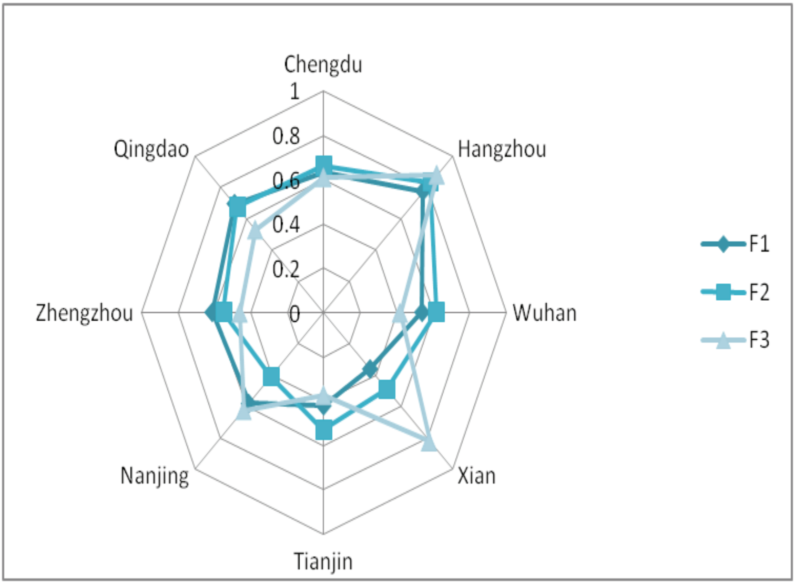

Fig 5-2 The investment scores in the second-tier city group

As can be seen from Figure 5-2, although the investment business of second-tier cities is not as good as that of first-tier cities, there are still some different investment characteristics in these cities. Hangzhou and Qingdao have relatively higher investment amount scores than the others, the fixed investment period score of Hangzhou is significantly better than those of the other cities, and Xi'an and Hangzhou have high reinvestment scores relative to the other cities. Therefore, it can be inferred that there is still much space for development of internet finance in China's second-tier cities.

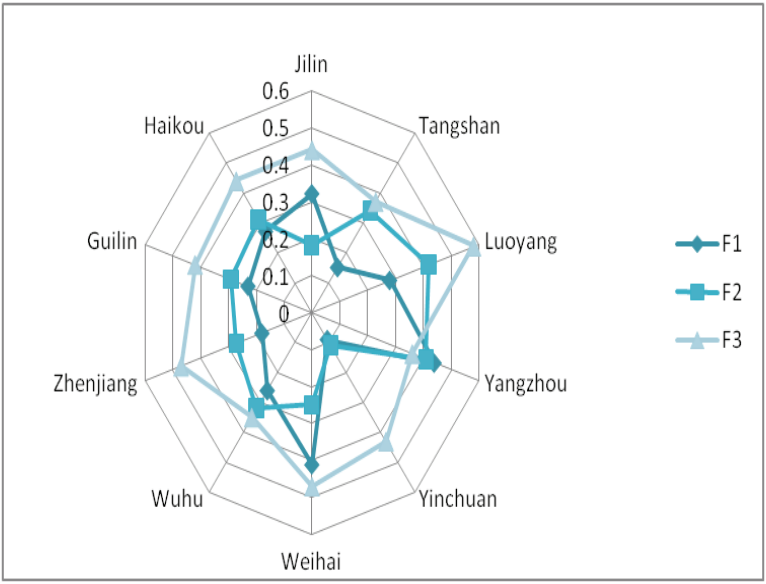

Fig 5-3 The investment scores in the third-tier city group

As can be seen from Figure 5-3, there is a big gap between the internet financial business of the third-tier cities and that of the first-tier and second-tier cities. Investment finance in third-tier cities is generally of a low amount and short term. Relatively speaking, the reinvestment score is relatively higher than the other two scores. Among them, Yangzhou and Weihai have relatively high scores in the amount of investment, Luoyang and Yangzhou have relatively high scores in the period of investment, and Luoyang and Weihai have relatively high scores in reinvestment. So, if a company wants to make a profit from internet finance, it will take some detailed analysis to choose the proper cities.

In summary, through the study of the investment situation, we can find that the investment variables are related to each other. Through factor analysis, we determined three main factors, which are mainly comprised of fixed investment amount, fixed investment period and revolving invest amount. With these factors, we explored the scores of different region groups; there are obvious regional differences in internet financial investment in China. Relatively speaking, in the first-tier cities such as Beijing and Shanghai, internet financial investment has been accepted by a vast number of consumers and has a wider customer acceptance, while in the second-tier and third-tier cities, internet finance has not yet been popularized and there is still a broad market space which will be expanded.

5.3 Factors for successful borrowing

Nowadays, more and more researchers put their focus on borrower features. Chen and Han (2012) conducted a comparative study of online P2P lending practices in the US and China and found that two categories of credit information, "hard" and "soft" information, may have profound influences on lending outcomes in both countries, but lenders in China are more reliant on "soft" information. Lin, Prabhala, and Viswanathan (2013) found that the online friendships of borrowers act as signals of credit quality. Zhang (2017) takes into account the effect of the guarantee mechanism on a loan's cash flow and calculated the expected internal rate of return of each loan from borrowers.

Also, there are some research discusses on the mechanism of P2P financing. Following the "all-or-nothing" rule, many loans fail due to insufficient pledges/money in their funding durations. Thus, automatically accessing and finding potential lenders early is crucial for loans. Zhang et al. (2018) proposed a hybrid random walk approach by combining both collaborative filtering and content-based filtering, which can be adapted to loans at any funding stage (e.g., the starting stage). Li, Wu, and Tang (2018) introduced a BP neural network interval estimation (BPIE) algorithm to predict the borrowers' borrowing limits and interest rates based on their characteristics and simultaneously developed a new parameter optimization algorithm (GBPO) based on the genetic algorithm. Song, Chen, Zhou, and Wu (2018) conducted an experiment with a novel two-stage slacks-based measure data envelopment analysis with a non-cooperative game; the performance efficiency of each stage as well as the comprehensive efficiency was evaluated to simulate the progress of P2P financing. Similarly, Liu, Qiao, Wang, and Li (2019) modelled a three-stage game to investigate optimal risk control ability and corresponding optimal prices of P2P lending platforms under different tariffs and agents' homing choices. Risk-price coefficients for lenders and borrowers were introduced to measure the impact of risk control ability on prices, where higher risk-price coefficients indicate that prices are more sensitive to risk control ability.

In this section, we put our focus on the factors of borrowers, and through the analysis of data coming from Credit Ease, we hope to find out the factors that affect the successful rate. And this section is separated into two parts: one is about the data and variables; the other is the empirical results.

5.3.1 Data and variables

The data in this research comes from Credit Ease in China. There are in total 769,365 detailed observations of borrowers during February 2008 to October 2011. Table 5-10 shows the details of the borrowers.

Table 5-10 The information variables of borrowers

Variable	Notes
borrow_id*key	
client_id*key	
acc_return	return date
borrow_purpose	1: consume;2: operation; 3: emergency; 4: other
apply_low_amount	
apply_high_amount	
borrow_sort	1:education; 2: house; 3: car; 4: enterprise; 5: other
gender	0: female; 1: male
degree	1: high school and below; 2: college; 3: graduate; 4: doctoral
birthday	
census_province	
census_city	
marital	1: married; 2: unmarried
child	
house	1: has house without loan; 2: has house with loan; 3: no house; 4: other
c_income	
credit	the total numbers of credit card
work_time	the working years

enterprise_type	1: public; 2. private; 3: other
ihas_local_phone	1: have; 2: do not have
ihas_mobile_phone	1: have; 2: do not have
ihas_email	1: have; 2: do not have
ihas_company	1: have; 2: do not have
ihas_company_phone	1: have; 2: do not have

We analyzed the binary variable "whether the loan is successful or not" as the dependent variable, and we used the binary logistic regression model expressed as $P = \frac{Exp(\beta_0+\beta_1 x_1+\cdots+\beta_m x_m)}{1+Exp(\beta_0+\beta_1 x_1+\cdots+\beta_m x_m)}$ to conduct the empirical analysis.

Firstly, we made a correlation analysis of the basic information variables of borrowers. The variables that influence the success of borrowing are marital status, children, degree, house, borrowing purpose, borrowing sort and income. The statistics for these variables are shown in Table 5-11.

Table 5-11 The statistics of independent variables

Marital	Number	Percentage	Grand total	Cumulative proportion
	1	0	1	0
0	2,526	1.42	2,527	1.42
1	102,733	57.72	105,260	59.14
2	57,929	32.55	163,189	91.69
3	13,825	7.77	177,014	99.46
4	956	0.54	177,970	100

Child	Number	Percentage	Grand total	Cumulative proportion
	1	0	1	0
0	2,805	1.58	2,806	1.58
1	103,799	58.32	106,605	59.9
2	71,365	40.1	177,970	100

Degree	Number	Percentage	Grand total	Cumulative proportion
	1	0	1	0
0	6,492	3.65	6,493	3.65
1	50,716	28.5	57,209	32.15
2	65,175	36.62	122,384	68.77
3	42,227	23.73	164,611	92.49
4	3,337	1.88	167,948	94.37
5	217	0.12	168,165	94.49
6	9,805	5.51	177,970	100

House	Number	Percentage	Grand total	Cumulative proportion
	1	0	1	0
0	3,560	2	3,561	2
1	60,592	34.05	64,153	36.05
2	49,622	27.88	113,775	63.93
3	48,520	27.26	162,295	91.19
4	13,169	7.4	175,464	98.59
5	2,506	1.41	177,970	100

Borrow purpose	Number	Percentage	Grand total	Cumulative proportion
1	73,597	41.35	73,597	41.35
2	48,746	27.39	122,343	68.74
3	35,894	20.17	158,237	88.91
4	13,540	7.61	171,777	96.52
5	6,193	3.48	177,970	100

Borrow sort	Number	Percentage	Grand total	Cumulative proportion
1	71,758	40.32	71,758	40.32
2	429	0.24	72,187	40.56
3	32,781	18.42	104,968	58.98
4	24,626	13.84	129,594	72.82
5	19,543	10.98	149,137	83.8
6	11,854	6.66	160,991	90.46
7	16,878	9.48	177,869	99.94
8	73	0.04	177,942	99.98
9	26	0.01	177,968	100
10	2	0	177,970	100

From Table 5-11, we know that the applicants are mainly between 20 and 50 years old, and 80% of them are male. The proportion of borrowers with medium educational background is high, accounting for more than 90%. There are fewer lower-educated and higher-educated borrowers on the application list, and there is little difference between genders. The applicants are mainly located in more than 20 provinces in East, South and North China. Due to the differences in law and economic development, there are few application data in Mongolia, Xinjiang, Tibet and other autonomous regions.

The average value of the loan applications is 4,400 RMB per month, which is higher than the mode of 1,400 RMB per month; the minimum value is 0 RMB, and the maximum value is 416,666 RMB. The main types of borrowing are "new salary loan" and "elite loan". Most of the applicants are middle class with fixed salaries. Their main borrowing purpose is consumption, accounting for more than 40%, less operation and emergency. Nearly 30% of the people are reluctant to disclose the purpose of borrowing.

5.3.2 Empirical results

The empirical results of the success rate are shown in Table 5-12.

Table 5-12 The empirical result of application successful rate

Variables	Coef.	S.E.	Wals
Borrow_purpose	-0.170***	0.01	320.943
Borrow_sort	-0.079***	0.021	14.027
Degree	0.186**	0.01	316.266
Marital	-0.099	0.013	55.852
Child	-0.042*	0.009	348.162
House	-0.262***	0.012	468.726
Loan_level	1.431	0.025	265.025
Income	0.016***	0.005	243.128
Credit	-0.081***	0.023	173.243
Age	0.0823	0.01	179.211
Constant	0.294**	0.091	10.484

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%, *** indicates that the p-value is less than 1%.

From Table 5-12, we can see that the Borrow_purpose, Borrow_sort, house, income and credit variables have obvious relationships with the success rate; degree can influence the success rate to a degree; and the child variable relates to the success rate to a significant degree.

For Borrow_purpose, Table 5-12 shows that borrowing for business and emergency purposes will reduce the success rate of borrowing. Operating risks, if not profitable, may not be able to recover the principal and interest after maturity; borrowers who borrow for emergency use may be forced by the situation, greatly increasing the motivation to conceal their true information, or even directly take out fraudulent loans, so the success rate of loan approval for emergency use is lower than others.

For Borrow_sort, Table 5-12 shows that borrowing for education will lift the success rate of borrowing. Educational lending has a positive effect on borrowers' ability. Usually, when borrowers have a better education experience, their income will improve, so the success rate of education is higher. And this result is consistent with the degree variable.

The coefficients of the house variables show that "mortgage with house" has a negative relationship with the success rate. This is meaningful, because "mortgage with house" is equivalent to the user who has passed the bank's initial qualification assessment, which is more authentic than the user information without a house.

The higher the credit amount is, the lower the success rate of borrowing. This may be because the credit information is not perfect in China, so the P2P platform cannot access the total credit balance of the applicant. Among 110,000 applicants, there are 105,500 people whose credit limits are less than 100 RMB, so the company does not know the credit line exactly, so they may take more consideration with those who have high credit amounts.

It is easy to understand the relationship between income, education, number of children and the success rate. High-income people are more likely to get loans. Likewise, highly-educated people tend to have relatively high incomes, so they are more likely to get loans. Generally speaking, the more children the family has, the higher the family's expenditure is, so there is always greater economic pressure with families with more children, and these families always have more difficulty obtaining loans.

China is a huge potential market for internet financing. Internet financing can be more effective than traditional borrowing, which matches the fund gap between the supply and demand sides. Through the analysis of Credit Ease's data, we can find that there are two key factors for the success of borrowing. One is the family's economic situation. When the family has a better income and lower expenditure, the success rate will be greatly improved. The second is the use of loans. When loans are used for business or education, the success rate will be lifted obviously. Therefore, we can conclude that internet financing is helpful to reduce the financial constraints of consumers to a certain degree, and if one can reasonably use this method to get credit, it may offer the chance to change one's life.

6. Credit Reference System

In 2013, China promulgated regulations on the management of the credit industry, which has sped up the track of the credit system. In 2014, China issued "the outline of credit system construction plan (2014–2020), which proposed to build the whole society credit system on the basis of information resources sharing". In 2015, eight institutions were permitted to conduct personal credit investigation services. By the end of June 2017, there were 92.6 million personal credit reports in the credit system. Between January and May 2017, the average daily query of personal credit report was 3.43 million times.

Now, the basic consumer data, mortgage data and credit data are becoming the main foundation for the credit reference, and the development of big data provides a possibility to explore deeper data application. With the increasing demands of credit information, user classification and user portrait are the key to the practical application of credit information. Therefore, it can be said that user classification is the first step in the process of gathering credit information, which restricts the application value of the whole information.

In this chapter, we first discuss the information query demand of the personal credit reference system in China, and then we use a clustering algorithm to portray the personal users in the system. Finally, we discuss the privacy protection of credit information.

6.1 Query demand

In the late 1980s, the personal credit reference system began to be established. After developing for more than 30 years, China has formed a multi-level credit information system based on central and local governments, which is controlled by the government and supplemented by market credit agencies.

However, the construction of China's credit system is facing a big contradiction. On the one hand, China's personal credit reference system is still immature, the coverage is relatively low and the information sharing mechanism is imperfect, and the credit service market is underdeveloped. On the other hand, with the development of China's economy, the credit information query demand has become more and more intense. Especially since the outbreak of internet financing in 2013, the financial industry has become more and more vocal in improving the personal credit reference system. Therefore, it is particularly important to strengthen the construction of the personal credit reference system.

6.1.1 Development

Reviewing the development history of China's credit industry over more than 30 years, it can be divided into four stages: budding stage, starting stage, initial development and rapid development.

For the budding stage, after the reform and opening up, due to the increase in China's foreign trade, the need to strengthen the credit status increased. At the same time, with the development of domestic bond and stock markets, the demand for professional assessment of corporate credit also grew. Under the joint stimulation of these needs, China's credit reporting industry has sprouted.

The starting stage is always seen as the early 1990s to 2003. In the early 1990s, a number of third-party credit investigation agencies appeared in China. Before the outbreak of the Asian financial crisis, many commercial banks began to use the bank credit data to establish a personal credit service system in order to strengthen the central credit system. In 1997, the People's Bank of China began to organize financial institutions to construct a credit reference system. At this stage, the overall development level of the domestic credit industry was relatively low; for example, the credit agencies often had government background, the data was incomplete, and the business scale was small, and it could not meet the corresponding requirements of the socialist market.

The government proposed to build a sound social credit system in the Tenth National People's Congress in 2003. Since then, the construction of China's credit system has entered a relatively rapid development. A major feature of this stage is that the government began to set up relevant departments to regulate and guide the development of China's credit reference system. The People's Bank of China established the Credit Administration and the Credit Center in 2003 and 2006, respectively. The Credit Administration is responsible for managing agencies, formulating regulations and developing plans for the industry. The Credit Center is mainly responsible for the construction of the credit reference system and a national unified credit database. In January 2006, the personal information basic database was officially put into operation. In the same year, the enterprise credit basic database received queries from outside.

After the financial crisis in 2008, the construction of China's credit reference system entered a stage of rapid development. After the crisis broke out, the Chinese government put great importance on the construction of the social credit system and continuously implemented relevant laws to regulate the credit market. In March 2013, the Regulations on the Management of the Credit Industry were officially implemented. In the same year, the first domestic credit industry development report was published. This report analyzed in detail the development history, opportunities and challenges faced by China's credit industry from 2003 to 2013. In October 2013, the Outline of the Credit System Construction Plan (2013–2020) was issued. The outline detailed plan gives a guideline for the development direction and states the major tasks of China's credit industry from 2013 to 2020.

At present, China's credit reference system has been basically completed. The unified personal and corporate credit system began to operate in 2006. As of October 2014, 19.63 million enterprises and 850 million natural persons were included. The system includes most of the bank credit information and a small amount of non-bank information such as financial data in financial companies, trust investment companies and leasing companies. The Credit Center is gradually introducing small credit institutions such as financing guarantee companies and microfinance companies to continuously improve the system.

In addition to the basic database of financial credit, the Credit Center has also established the Network Financial Credit System (NFCS). The system mainly provides information sharing and query services for online loan companies which engage in P2P services. As of December 2014, the system had access to 370 online lending institutions, with a total of 524,000 customers.

In terms of credit services, credit report enquiries are at the core. At present, the personal credit report from the People's Bank of China in the credit reference system has reached an average of one million inquiries per day, and the enterprise credit report has reached an average over 270,000inquiries per day. Figure 6-1 shows the annual query amount for the credit reference database; the number of personal inquiries is marked in blue, and the number of corporate inquiries is marked in red. The credit information must be widely used in the credit risk evaluation, which is providing huge amounts of information for individuals and businesses in financing.

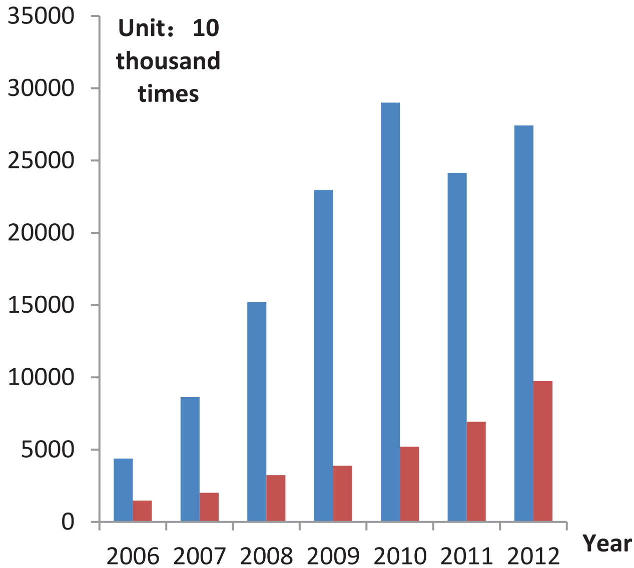

Fig 6-1 The number of annual queries from the credit reference system

In addition to the People's Bank Credit Center, third-party credit institutions in the domestic credit market also play an important role. According to the China Credit Industry Development Report, as of 2012, there were more than 150 credit agencies in China. The main business of these institutions is mostly directed at corporate credits such as corporate credit ratings. The market shares of these institutions can be found in Figure 6-2.

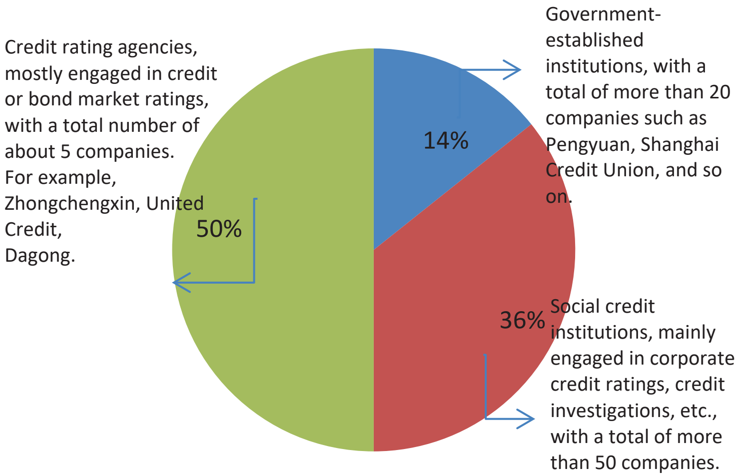

Fig 6-2 The market shares of credit agencies

In the past two years, big data credits developed through the internet have become the leader in the credit industry. In January 2015, the central bank issued business licenses to eight institutions, including Tencent Credit and Sesame Credit. This is a signal that the government allows the internet social platforms to participate in the personal credit industry. These internet companies have unique advantages in the credit dimension, and their massive back-end data are a huge precious resource for the personal credit industry. Take Tencent as an example: In 2014, Tencent had 800 million QQ accounts and more than 500 million WeChat accounts.[4] Now, the key issue is how to use data mining and corresponding data integration technology to measure personal credit status, to form a unified core database, and to expand the amount of personal credit information evaluations and then improve China's personal credit industry.

At present, the domestic credit reference system cannot meet the market's credit demand, and the gap between supply and demand is large. For example, only the People's Bank Credit Information Center is responsible for personal credit in China, which has a low personal

[4] https://wenku.baidu.com/view/f467296558eef8c75fbfc77da26925c52cc5912c.html. Published date: 2018-08-09.

credit coverage rate.

6.1.2 Query demand

We used the total social credit amount and personal credit system coverage to build a personal credit supply-demand model and analyze the supply and demand of the individual credit reference system.

In order to quantify the supply and demand of credit information, we tried to build a regression analysis. The dependent variable is the personal credit demand. Considering that the credit data query is one of the market's reflections on the credit supply, we transformed it using the annual query volume (query) of the personal credit reference system.

One explanatory variable of the model concerns the information degree of the personal credit reference system, which is expressed by the coverage rate (coverage, coverage rate is the number of people included in the credit reference system dividend by the number of people over 18 years old in China). Another explanatory variable of the model is the total amount of personal loans, which is the social loans (loan) declared by banks.

The supply model and demand model are built as follows (6.1).

$$Query = \beta_1 Coverage + \beta_2 Loans + \varepsilon \qquad (6.1)$$

The data is shown in Table 6-1. And the result can be found in Table 6-2.

Table 6-1 The data of the credit reference system, using a model

Year	Query	Coverage	Loan
2006	0.4380	0.514	3.15
2007	0.8627	0.564	3.63
2008	1.5205	0.597	4.90
2009	2.2966	0.618	9.59
2010	2.9004	0.714	7.95
2011	2.4146	0.720	7.47
2012	2.7427	0.731	8.20
2013	3.4000	0.746	8.89
2014	3.9240	0.759	9.78

Table 6-2 The results of regression

	df	Sum Squares	Mean Square	F	P
Regression	2	10.1173	5.0587	69.6617	0.0001
Residual	6	0.4357	0.0726		
Sum	8	10.553			
Root MSE	0.2695			R Square	0.9587
Dependent Mean	2.2777			Adj R-Sq	0.945

Variables	Coef.	S.E.	T	P
Constant	-4.1028***	0.9023	-4.5471	0.0039
Coverage	7.5431***	1.8889	3.9934	0.0072
Loan	0.1958***	0.0673	2.9082	0.027

Note:* indicates that the p-value is less than 10%, ** indicates that the p-value is less than 5%,*** indicates that the p-value is less than 1%.

From the results of Table 6-2, we can see that credit coverage and total social loans are positively correlated with credit query demand. Among them, for every increase in credit coverage, the demand for credit inquiries will expand by more than 7%, while for every increase in the total number of social loans, the demand for credit inquiries will expand by about 0.2%. According to the survey of the People's Bank, at the end of 2017 there were over 100 million persons who had records in the credit reference system, and that number will double in the next five years. Even if the total number of social loans does not increase, the demand for credit enquiries will increase by 7 million; this is a real change for the credit reference system.

Therefore, the Credit Center needs to substantially increase the collection of personal credit and include more natural persons; otherwise, it still has a deep gap between supply and demand. In summary, it can be seen that increasing demand for credit information collection, improving the information sharing service and providing more abundant credit products are the most urgent tasks for the credit reference system.

6.2 User portrait

Current research on the personal credit reference system is relatively scarce. Many previous research efforts focused on the description of consumer credit behavior, including studies by Diallo and Al-Titi (2017) and Nakamura and Roszbach (2017). According to the findings of Mikhed and Vogan (2018), a clustering algorithm is often used to analyze consumer characteristics.

Clustering algorithms can be divided into four categories: hierarchical clustering, segmented clustering, mixed clustering and grid clustering, as discussed by Anderberg (1973), who found k-means hierarchical clustering to be the most widely used algorithm. Current research focuses on the selection of initial settings, such as in studies by Huang et al. (2014), Kumar and Reddy (2017) and Zahra et al. (2015), and algorithm performance enhancements such as in studies by López-Rubio, Palomo, and Ortega-Zamorano (2018), Liu, Chang, and Li (2013), and Lu, Qin, Cao, Liu, and Wang (2014). However, according to the research of Wu, Lin, Zhao, and Yan (2018), the k-means algorithm always deals with numeric attributes, and for frequency attributes this algorithm is tested to be invalid, greatly limiting its application. Datta, Bhattacharjee, and Das (2018) modified the traditional k-means clustering algorithm and the standard hierarchical agglomerative clustering algorithms to make them directly applicable to datasets with missing features. They also utilized a penalized dissimilarity measure, which is referred to as the feature-weighted penalty-based dissimilarity. Ting et al. (2018) proposed the use of mass-based dissimilarity, which employs estimates of the probability mass to measure dissimilarity, to replace the distance metric. They showed that we can deal with special data by adjusting to different measures. Also, some existing works observe that most of the existing clustering algorithms attain good performance for specific problems but are not robust enough for a wide range of data analysis problems, they and suggest the combination of multiple settings. Related work can be found in the research of Hu and Wong (2013) and López-Rubio, Palomo, and Ortega-Zamorano (2018).

6.2.1 Method

In k-means, we are using the number k to partition into k clusters, and the way we do the clustering is using the metrics called means or average, in a numeric dimension or centroid in 2 or more dimensions. The k-means algorithm is also called the k-centroid algorithm in some cases. So, we start with n objects or instances and the value of k. The outcome of the k-means algorithm is to find a set of k clusters such that all the instances in each cluster are bunched together. That is, they are tightly bound into clusters using what is called the squared error criterion. The ultimate goal of the algorithm is to find out k clusters which can minimize the sum of squares within clusters as in (6.2)

$$argmin \sum_{i=1}^{k} \sum_{x \in S_i} \|x - u_i\|^2 \tag{6.2}$$

where u_i is in the mean of samples in the same cluster.

If there are k means which are listed as $m_1, m_2, \cdots m_k$, the algorithm can be done as follows: Firstly, each observation is allocated to the cluster to minimize the sum of squares within the group, which can be defined as $S_i = \left\{x_p : \left\|x_p - m_i{}^{(t)}\right\|^2 \leq \left\|x_p - m_j{}^{(t)}\right\|^2, \forall j, 1 \leq j \leq k\right\}$, and then we make the changes to the mean of new clusters as $m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|} \sum_{x_j \in S_i^{(t)}}^{x_j} x_j$. Repeat these two steps until the intra-cluster mean does not change. The steps can be listed as follows.

Step 1: Data standardization

For$X = \{x_1, x_2, \ldots, x_n\}$, we calculated the mean and variance, and then transformed each element to $x_{ij}{}' = \frac{x_{ij} - \bar{x}_j}{\sigma_j}$, with $\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$, $\sigma_j^2 = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$, $i = 1,2, \ldots, n$, $j = 1,2, \ldots, m$, thus obtaining the ordered standard set $\{x_i\}$.

Step 2: Construct the distance matrix

We used Euclidean distance to calculate the similarity between the two sets of $\{x_i\}$ and then calculated the mean of the initial clusters.

Step 3: Clustering

For a new sample, calculate the distance between every cluster means, and choose the min one cluster, then add the sample into the cluster. When all the samples are in the cluster, then calculate the new means of the clusters. Repeat this process until all the samples are in clusters, and the means of clusters will not change.

6.2.2 Data and variables

In this experiment, we used a credit card database as the sample, as provided by the personal credit reference system in China. The data set spans from 2004 to 2009, with a total number of 65,536 accounts and 31 variables.

Firstly, we explored the basic information of the accounts. In this sample, there are more than 40% of users with a bachelor's degree or above. Nearly 80% of users are married. The geographic distribution of accounts is shown in Figure 6-3.

From Figure 6-3, it can be seen that users in the personal credit reference system are mainly from Beijing, Jiangsu, Shanghai, Shandong, Zhejiang and Guangdong. All of these regions in China are developed, and we can state that the regional distribution of this sample is reasonable and representative of all users.

Next, we examined the development of accounts. We examined trends in the ratio of debit accounts, the ratio of credit accounts and the number of effective accounts, as shown in Figure 6-4.

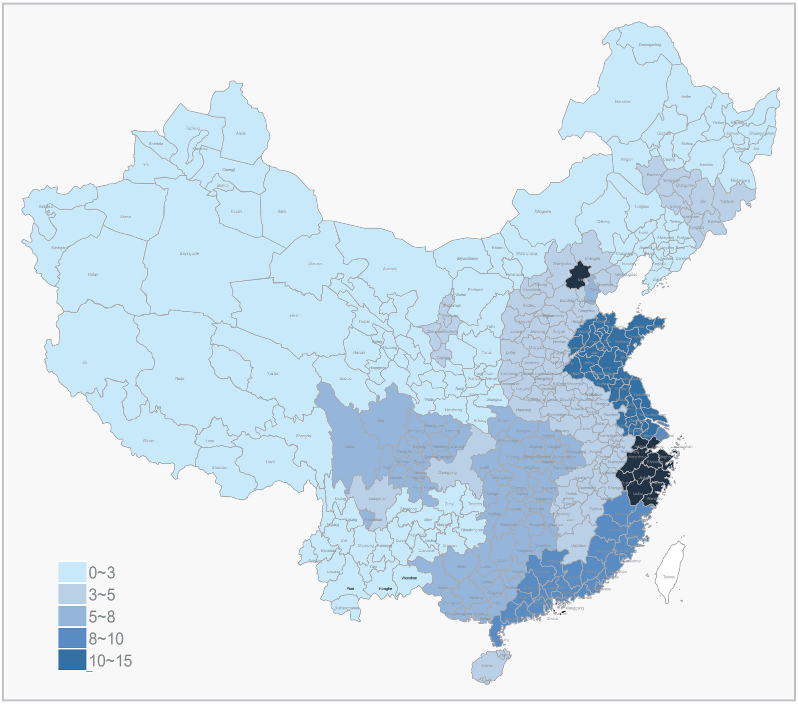

Fig 6-3 The geographic distribution of accounts

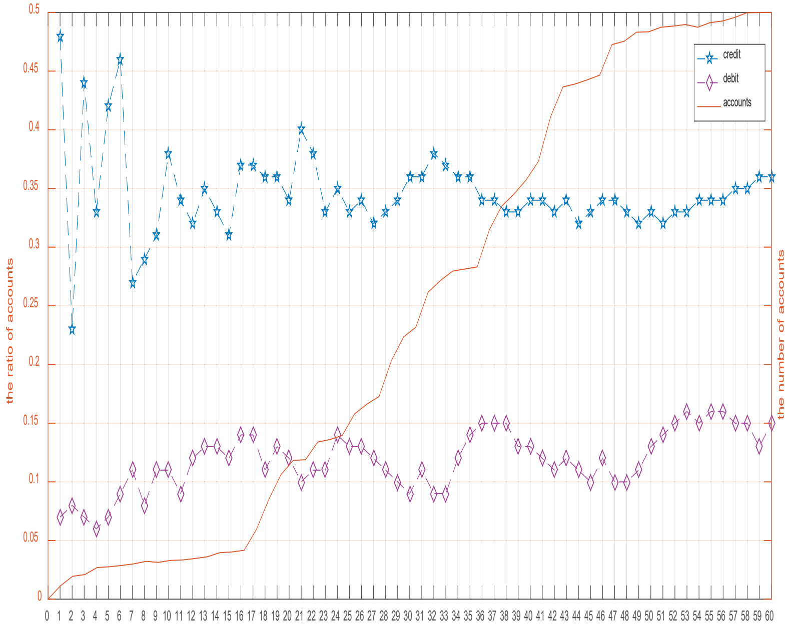

Fig 6-4 The trend of accounts and the ratio of credit/debit

From Figure 6-4, we can see that the number of accounts showed rapid growth with an almost exponential trenduntil2008. The growth then slowed, and the trend approached stability. The ratio of credit accounts changed rapidly initially, and then it varied at a level between 30% and 35%. The ratio of debit accounts changed relatively less than the ratio of credit accounts, and it has been around 10% and never over 16% in last five years. Therefore, it can be said that the users in this system have become stable, and we can use these samples to perform clustering.

Furthermore, in order to check whether there were any great changes in the sampled accounts, we took average annual income and average total debt to explore the trend of economic attributes in individual accounts, as shown in Figure 6-5.

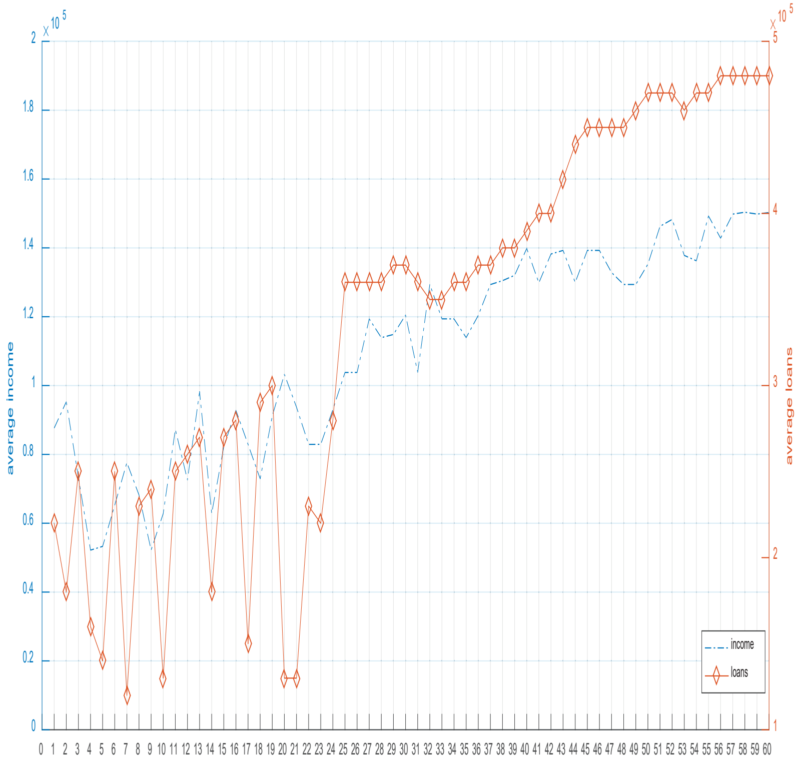

Fig 6-5 The trend of average annual income and average total debt

From Figure 6-5, we can see that between 2004 and 2005 the average total consumer debt changed rapidly. After 2006, the average total user debt growth became relatively stable. Meanwhile, it can be seen that the average annual income increased, almost following a line, and its speed of growth was much less than that of the average total debt. Nonetheless, it can be inferred that average individual annual income and average individual total debt have become stable in recent years, and these samples can be used to represent all consumers in the system, and so the clusters are reasonable and stable. The sample for analysis comprises consumers in the personal credit reference system with dates after 2006, and there are 53,892 such records.

In order to make uniform the distance of the variables, we segmented the continuous variables, and the details are shown in Table

6-3. The observations marked with blue are the variables that were deleted for a certain reason.

Table 6-3 Variables and data preprocessing

Variables	Data preprocessing
ID number	Delete because of privacy
Name	Delete because of privacy
Query time	Only used with year
The number of queries	
Gender	1 for male, 2 for female
Age	1 for 20–30; 2 for 31–40; 3for 41–50; 4 for 51–60; 5 for over 60
Marital status	1 for married; 2 for unmarried
Telephone	Delete because of privacy
Affiliation	Delete because of privacy
Work phone	Delete because of too many missing values
Home phone	Delete because of too many missing values
Work address	Delete because of privacy
Home address	1 for home address matches mailing address; 2 for other
Education degree	1 for without schooling; 2 for junior high school and below; 3 for high school and secondary school; 4 for undergraduate or college; 5 for master; 6 for doctor
Region	Zip code
Place of birth	Zip code
Annual income	1 for 10,000 yuan and below; 2 for 10,000-50,000 yuan; 3 for 60,000-100,000 yuan; 4 for 110,000-200,000 yuan; 5 for 210,000-500,000 yuan; 6 for 510,000-1,000,000 yuan; 7 for more than 1,000,000 yuan
Industry	Following GB/T 4754-2017; total 21 kinds
Occupation	Following GB6565-86; total 8 kinds
Number of loans	

Date first acquired loan	Only used with year
Total debt	1 for 10 wan yuan and below; 2 for 10–50 wan yuan; 3 for 51–100 wan yuan; 4 for 101–300 wan yuan; 5 for 301–500 wan yuan; 6 for 501–1,000 wan yuan; 7 for more than 1,000 wan yuan
Debt balance	1 for 10 wan yuan and below; 2 for 10–50 wan yuan; 3 for 51–100 wan yuan; 4 for 101–300 wan yuan; 5 for 301–500 wan yuan; 6 for 501–1,000 wan yuan; 7 for more than 1,000 wan yuan
Repayment (monthly)	1 for 1 wan yuan and below; 2 for 1.1–2.0 wan yuan; 3 for 2.1–5.0 wan yuan; 4 for 5.1–10 wan yuan; 5 for more than 10 wan yuan
Number of credit cards	
Date first acquired credit card	Only used with year
Credit card amount	1 for 1 wan yuan and below; 2 for 3 wan yuan; 3 for 3–5 wan yuan; 4 for 5–10 wan yuan; 5 for 10–30 wan yuan; 6 for more than 30 wan yuan
Credit balance	1 for 1,000 yuan and below; 2 for 1,001–30,00 yuan; 3 for 3,001–5,000 yuan; 4 for 5,001–10,000 yuan; 5 for more than 1 wan yuan
Debit balance	1 for 100 yuan and below; 2 for 101–500 yuan; 3 for 501–1,000 yuan; 4 for 1,001–3,000 yuan; 5 for more than 3,000 yuan
The number of guarantees	

Then, we used maximum/minimum-normalized processing to put all of these records into an interval [0, 1] in order to calculate the distance matrix of the sample, to give the distance the same dimension.

6.2.3 Dynamic clustering

According to the clustering algorithm, we explored the orthogonal process to find the characteristic elements in 31 variables, as described in step 2 of the algorithm. It was found that there were 10 variables in the results, which were: gender and marital status (x1), the number of queries (x2), education degree (x3), industry (x4), the number of guarantees(x5), the number of credit cards (x6), the number of loans (x7), credit balances (x8), total debt (x9)and annual income (x10), listed by their variance.

With the 53,892 records and initial 10 variables, we used the maximal number of occurrences to determine the initial number of clusters and used the mode of the final cluster to determine the initial record core.

We explored the results by constructing a diagram to show the relationship between cluster number and distance as shown in Figure 6-6.

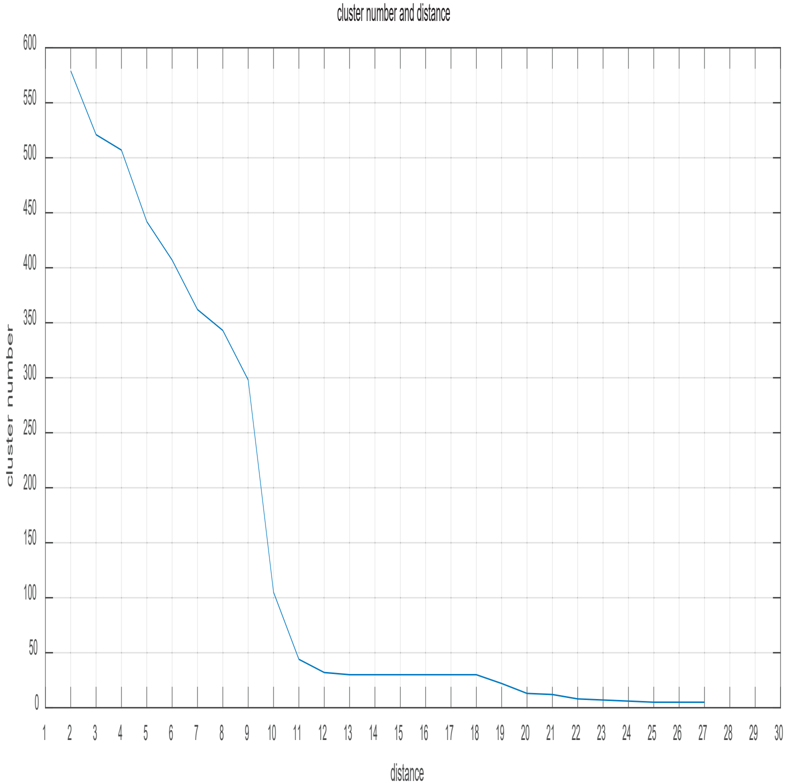

Fig 6-6 The number of clusters and the distance

From Figure 6-6, it can be seen that for distances between 13 and 18, the number of clusters is always 30, which is most representative of the experimental results and is the flattest area in the figure. Therefore, we set the number of initial clusters to 30.

The clustering growth process was performed using MATLAB 2016a, with the 30 initial classes. The clustering process is shown in Figure 6-7.

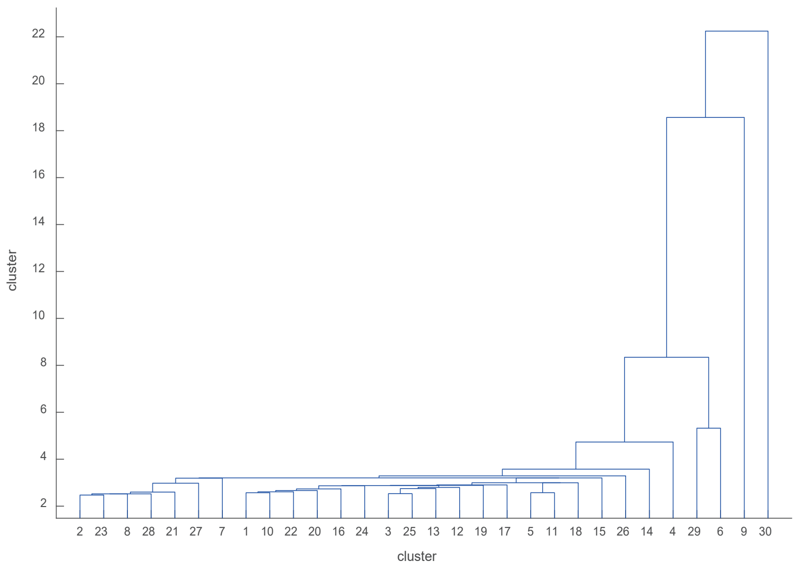

Fig 6-7 The combination process of clusters

As shown in Figure 6-7, there are five prominent peaks in the graph, and so we inferred that the number of final clusters was five. Furthermore, we summarized the growth process as shown in Table 6-4.

Table 6-4 The summary of dynamic clustering

Category	Records	Max distance within category	Nearest category
1	11,483	1.0926	3
2	8,682	1.0451	1
3	10,720	1.0218	1
4	2,595	1.8278	5
5	20,412	1.0917	4

The records in different clusters were not equal. Category 5 had the most records, and category 4 had the least. The maximum distance within categories was almost equal, and so it can be said that it was concentrated within classes. Category 4 had the largest distance, which was evidently different from the others. According to the distance by category, we found that categories 2 and 3 had the tendency to combine with category 1, and category 4 had the tendency to combine with category 5. We also explored the modes of the 10 core variables within these categories, and the results are listed in the Table 6-5.

Table 6-5 The modes of 10 variables with 5 categories

Mode	1	2	3	4	5
Gender and marital status (x1)	11	21	11	12	11
Number of queries (x2)	3	4	5	5	7
Education degree (x3)	4	3	4	3	5
Industry (x4)	3	5	4	9	10
Number of guarantees (x5)	3	1	2	2	3
Number of credit cards (x6)	1	0	3	3	5
Number of loans (x7)	1	1	3	2	4
Credit balances (x8)	2	1	3	4	4
Total debt (x9)	2	2	4	4	5
Annual income (x10)	3	2	3	4	5

Using the χ^2 test, it was found that 8 of the 10 variables were different at the 95% significance level for category 1 and category 5; gender and marital status (x1) and the number of guarantees (x5) failed the test. Five of the 10 variables were different at the 95% significance level for category 1 and category 3; gender and marital status (x1), education degree (x3), industry (x4), total debt (x9) and annual income (x10) failed the test. Seven of the 10 variables were different at the 95% significance level for category 3 and category 5; gender and marital status (x1), the number of credit cards (x6) and total debt (x9) failed the test.

6.2.4 Portrait

With the results of clustering, we selected the main three categories of users, which are category 1, category 3 and category 5, to explore detailed personas.

In order to explore the personas, we used distribution analysis of the most widely different variables to make a detailed comparison of the characteristics of these three categories.

For category 1 and category 5, we only explored the variables with the difference under significance level 0.01. There are four such variables, which are: the number of queries (x2), industry (x4), the number of loans (x7) and annual income (x10). Because the numbers of samples in the two categories are different, we used the percentage to describe the distribution within the variables; the results are shown in Figure 6-8.

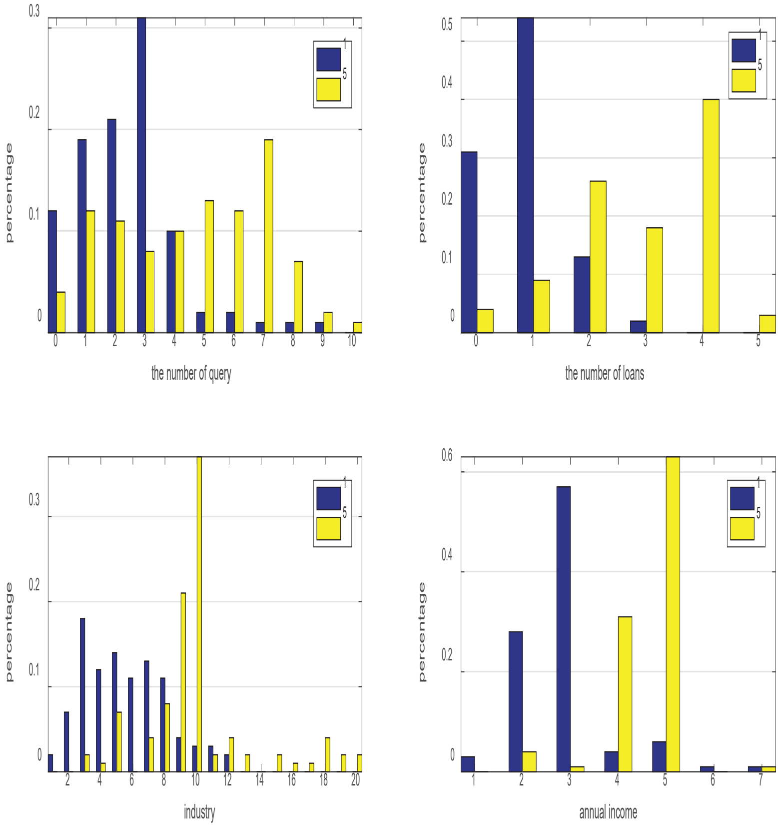

Fig 6-8 Comparisons between category 1 and category 5

As can be seen from Figure 6-8, with the number of queries (x2), in category 1 there are mostly less than 5, in category 5 there are more than 5, and the distribution in category 1 is more left skewed than in category 5.With industry (x4), users in category 1 are mainly in jobs in traditional industries, the extreme majority of users in category 5 are in jobs in information technology and finance. With the number of loans (x7), most users in category 1 have no loans or only took a loan once and no more than 3 times, most users in category 5 have more than twice taken a loan, and the distribution in category 5 is more inhomogeneous and right skewed than in category 1. With the annual income (x10), users in category 1 earn less money than users in category 5. To sum up, we can infer that users in category 1 are those who work in the traditional industry, are less wealthy and have less credit use; in contrast, users in category 5 are those who work in the modern developed industries such as financing, are wealthier and use more credit.

For category 1 and category 3, just as in the section above, we only explored the variables with the difference under significance level 0.01. There are two such variables, which are: the number of credit cards (x6) and credit balance (x8). Because the numbers of samples in the two categories are different, we used the percentage to describe the distribution within the variables; the results are shown in Figure 6-9.

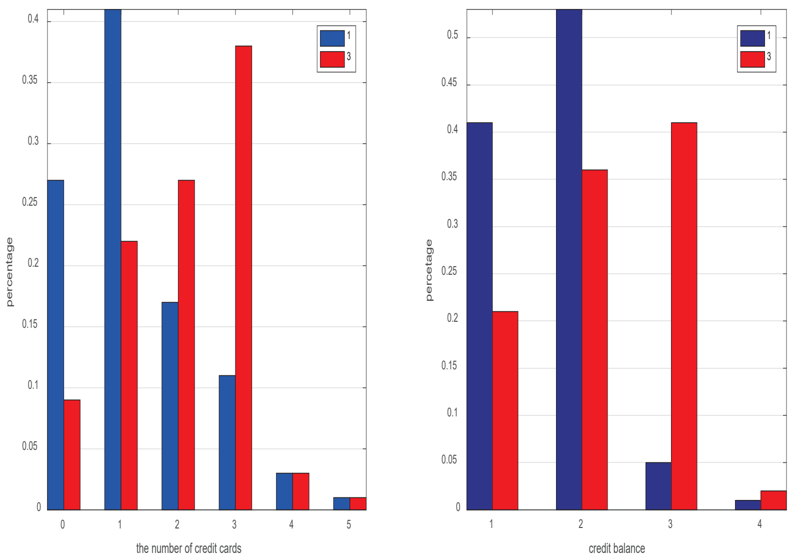

Fig 6-9 Comparisons between category 1 and category 3

As can be seen from Figure 6-9, with the number of credit cards (x6), users in category 3 have more credit cards than users in category 1, and the distribution in category 1 is left skewed compared with category 3. With credit balance (x8), most users in category 1 have less debt than users in category 3, and the distribution is relatively concentrated. To sum up, we can infer that users in category 1 are those ones who are inactive in the credit market, and they may be those who are relatively averse to taking risks; users in category 3 have more desire to use credit. So we can generalize that category 1 is conservative and category 3 is relatively active.

For category 3 and category 5, just as in the section above, we only explored the variables with the difference under significance level 0.01. There are three such variables, which are: industry (x4), credit balance (x8) and annual income (x10). Because the numbers of samples in the two categories are different, we used the percentage to describe the distribution within the variables; the results are shown in Figure 6-10.

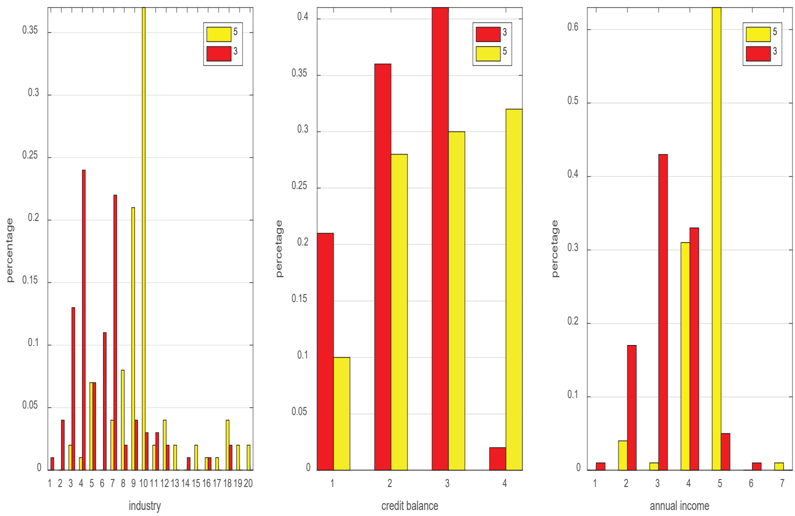

Fig 6-10 Comparisons between category 3 and category 5

As can be seen from Figure 6-10, with industry (x4), users in category 1 are mainly in jobs in traditional industries, and the extreme majority of users in category 5 are in jobs in information technology and finance. With credit balance (x8), most users in category 3 have lower credit balances than users in category 5, and the distribution is relatively concentrated. With annual income (x10), users in category 5 earn more money than users in category 3; most annual incomes in category 5 exceed 20 wan RMB. To sum up, we can infer that users in category 3 are those who have middle incomes, and users in category 5 are those who have high incomes and high level of debts.

Through the experiments, it can be found that the credit users can be divided into five categories. The three main categories can be summarized as: middle-income conservatives (category 1), middle-income activists (category 3) and high-income high lever activists (category 5). Furthermore, it can be seen that in the personal credit reference system, the portrait of the users can be drawn from the two basic dimensions, which are: income and credit. To sum up, user classification is the first step of credit information extraction, which restricts the application of the personal credit reference system.

6.3 Privacy of credit data

At present, China's credit system has been basically completed. This section mainly introduces the construction achievements of China's credit system from three aspects: legal system construction, credit information service and credit agencies. And then we provide a discussion on the protection of privacy.

The most important part of the existing legal documents in China's credit industry is the Regulations on the Administration of Credit Information. The regulations were officially implemented in March 2013. They mainly specified the rules for the industry construction, the regulatory agencies, the operating rules and the rights of use, which provided basic laws for the standardized development of the credit industry. The other two important documents are "Basic data guidance for the construction of the SME credit system" and "Basic Assumption Data Guidance for the Rural Credit System". These documents are an important legal basis for promoting the construction of credit systems in China's SMEs and rural areas. They are mainly used to guide how to better integrate rural credit and SME credit into the national credit system. At present, China's credit reference system has been basically completed and the law has formed a framework. It is reported that the annual income of the credit industry is more than 2 billion RMB.[5]

For the credit information service, the unified personal and corporate credit system began to operate in China in 2006. As of December 2014, 19.63 million enterprises and institutions and 850

[5] China Credit Industry Development Report.
http://upload.xinhua08.com/2013/1212/1386812896462.pdf
Published date: 2013-12.

million natural persons were included. The system information includes most of the bank credit information and a small amount of non-bank information such as information about financial companies, trusts, investment companies and leasing companies. The Credit Center is in charge of the database, and it provides the information service as queries and printing reports. In addition to the basic database of the credit reference system, the Credit Center has also established the NFCS. The system mainly provides information sharing and query services for internet financing companies which engage in financing activities such as P2P services. As of December 2014, the system had access to 370 internet financing companies, with a total of 524,000 customers.[6]

By the end of 2012, there were more than 50 credit agencies with social backgrounds, including the China Credit Reporting Company, United Credit Company, Dagong International Company and Shanghai New Century. Eight companies have individual credit investigating licenses: Sesame Credit, Tencent Credit, Qianhai Credit, Pengyuan Credit, Zhongchengxin Credit, Chi-Cheng, Koala Credit and Huadao Credit. These credit institutions are mainly engaged in credit investigation, rating analysis and guaranteeing business.

China's credit industry started in the mid-to-late 1980s, but the Regulations on the Credit Industry Management was introduced in 2013. So far, there is no detailed legal document in China to regulate the collection and use of credit information in the credit industry. The Regulations on the Credit Industry Management only played an indicative role, and many details were not clear enough. Though local government has also issued relevant regulations such as "Trial Measures for the Management of Shanghai Personal Credit Management", but the detailed requirements vary from place to place, resulting in irregularities in cross-regional information collection activities. In terms of the protection of personal privacy, laws such as the Personal Information Protection Act, the Commercial Protection Act and the Data Protection Act have not yet been implemented.

The private credit market system has not yet been formed. Among the credit agencies, the number of those who engaged in personal credit investigation is small, and they have few credit products and services. Only eight institutions of the central bank issued a personal

6 https://wenku.baidu.com/view/d1570b3d590216fc700abb68a98271fe910eaf8f.html?rec_flag=default&sxts=1564968538109 Published date: 2016-11-06.

credit business license in January 2015, and there is still a large gap between market needs and institution supply.

Credit data mainly include credit transaction activity records and credit status records. Specifically, personal credit data includes basic personal information (such as personal identity, occupation, taxation, court records, and so on) and credit transaction information (such as bank balance, guarantee, and so on). Corporate credit data includes basic corporate information (such as corporate legal person, registered capital, and so on), corporate credit information (such as liabilities, mortgage, and so on) and public information (such as corporate tax arrears, court enforcement, and so on). Credit agencies collect credit information about enterprises and individuals from financial institutions such as commercial banks and insurance companies, as well as non-financial institutions such as courts, industry and commerce bureaus.

After processing, the credit databases are formed. When the credit agency provides the credit service, it can extract the corresponding information from the credit reference database to form an objective and fair credit report. Some credit agencies will also conduct credit evaluations on the information subject based on collected credit data, such as the credit rating of the enterprise.

Due to the short development time and the delay in legal system construction, there are still many problems with China's credit reference system. The main problems may be the imperfect legal system and limited private credit information sharing. There are three processes with credit information sharing, which are: information collection, storage and information usage. Figure 6-11 shows the security issues that may arise in each process.

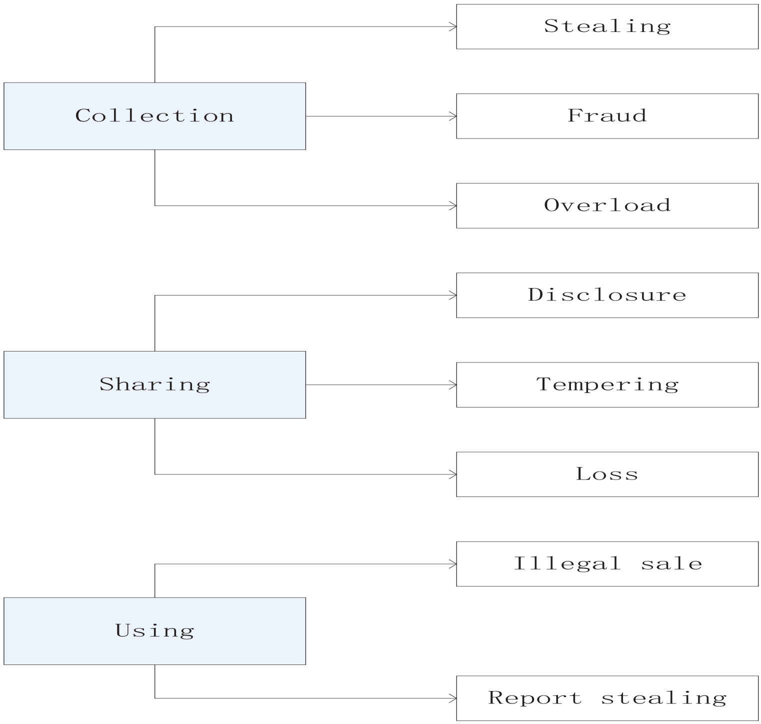

Fig 6-11 Main problems with credit information sharing

Privacy in information storage can be controlled by database technology and log management. Privacy issues in information sharing and usage need more laws to regulate and improve. On the one hand, the occurrence of some incidents has infringed on the legitimate rights and interests of persons, making more individuals unwilling to grant credit and unwilling to share data. On the other hand, the occurrence of some incidents also reflects the serious situation of information security and privacy protection for individuals and enterprises. At present, there are still no clear legal provisions on how to protect personal information and how to punish people who divulge information. The legal system for credit reference is still not perfect. Under these circumstances, China's credit data security mechanism urgently needs to be built.

A sound information sharing mechanism can balance the openness and security, and it can enable industry commerce, taxation, courts and different administrative regions across the country to share their database information safely and efficiently. Improving the information sharing mechanism can start from the following four points: Firstly, there is a need to improve the legal system for information sharing and enable legal supports in information exchanges between different departments and regions; the People's Bank is responsible for unifying the construction standards of credit reference systems based on various industries and regions, and now it needs to reduce the dimensions and build a public service sharing platform. Secondly, there is a need to improve the sharing incentive mechanism so that it can enable both parties to reach agreements on long-term interests. Thirdly, there is a need to protect the security of sharing information and set more laws, and we should learn experience from developed countries. Finally China must explicitly prohibit other organizations from sharing credit information data.

To sum up, through the supply-demand modeling analysis of personal credit information and the discussion of credit data security, we can reveal the conflict between credit data opening and credit data protection in China's credit reference system, and we also find that the legal system and the supervision mechanism are not perfect, which is the underlying cause of conflict. In order to resolve the conflict, we put forward four suggestions for the construction of China's credit reference system, including: strengthening the legislative protection of credit information, improving industry supervision, enhancing data management of credit agencies and reinforcing the information sharing mechanism.

References

Abdul Muhmin, Alhassan G. 2008. Consumer attitudes towards debt in an islamic country: managing a conflict between religious tradition and modernity? *International Journal of Consumer Studies* 32 (3):194-203.

Aghion, Philippe, George-Marios Angeletos, Abhijit Banerjee, and Kalina Manova. 2010. Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics* 57 (3):246-265.

Alonso, Cristian. 2018. Hard vs. soft financial constraints: Implications for the effects of a credit crunch. *Journal of Macroeconomics* 58:198-223.

Anderberg, Michael R. 1973. Chapter 6 - Hierarchical Clustering Methods. In *Cluster Analysis for Applications*: Academic Press.

Ando, Albert, and Franco Modigliani. 1963. The "Life Cycle" Hypothesis of Saving: Aggregate Implications and Tests. *The American Economic Review* 53 (1):55-84.

Auh, Seigyoung, and Chuan-Fong Shih. 2005. Balancing Giving-Up Vs. Taking-In: Does the Pattern of Payments and Benefits Matter to Customers in a Financing Decision Context? *Advances in Consumer Research* 33:139.

Awanis, Sandra, and Charles Chi Cui. 2014. Consumer susceptibility to credit card misuse and indebtedness. *Asia Pacific Journal of Marketing and Logistics* 26 (3):408-429.

Banerjee, Tanmoyee, Chandralekha Ghosh, and Malabika Roy. 2010. Borrowers in a village economy: An analysis of credit contracts across rural households. *Asia-Pacific Social Science Review* 10 (1):85-101.

Banerji, Sanjay, Rajesh S. N. Raj, and Kunal Sen. 2016. Monitoring Costs, Credit Constraints and Entrepreneurship. *The Manchester School* 84 (5):573-599.

Bauman, Zygmunt. 2001. Consuming Life. *Journal of Consumer Culture* 1 (1):9-29.

Bechlioulis, Alexandros P., and Sophocles N. Brissimis. 2019. Consumer debt non-payment and the borrowing constraint: Implications for consumer behavior. *Journal of Banking and Finance* 101: 161-172.

Becker, Thomas A., and Reza Shabani. 2010. Outstanding Debt and the Household Portfolio. *The Review of Financial Studies* 23 (7):2900-2934.

Brown, Jane, and Helen Woodruffe-Burton. 2015. Exploring emotions and irrationality in attitudes towards consumer indebtedness: Individual perspectives of UK payday loan consumption. *Journal of Financial Services Marketing* 20 (2):107-121.

Cai, Dongliang, Quanyun Song, Shuang Ma, Yang Dong, and Qiuhua Xu. 2018. The relationship between credit constraints and household entrepreneurship in China. *International Review of Economics & Finance* 58:246-258.

Caldieraro, Fabio, Jonathan Z. Zhang, Marcus Cunha, and Jeffrey D. Shulman. 2018. Strategic Information Transmission in Peer-to-Peer Lending Markets. *Journal of Marketing* 82 (2):42-63.

Chahal, Hardeep, Gurjeet Kaur, and Anu Rani. 2015. Exploring the dimensions of customer experience and its impact on word-of-mouth: a study of credit cards. *Journal of Services Research* 15 (2):7.

Chang, Chun-Wei, and Olumide Ijose. 2016. Measuring Customer Lifetime Value: An Application in Credit Card Industry. *Academy of Business Research Journal* 1:7.

Chen, Dongyu, and Chaodong Han. 2012. A comparative study of online P2P lending in the USA and China. *Journal of Internet Banking and Commerce* 17 (2):1-15.

Chen, Tser-Yieth, and Fang-Ju Lin. 2011. Using Relationship Norms to Create Appropriate Relationship Value: Evidence From the Credit Card Industry. *Journal of Relationship Marketing* 10 (1):28-42.

Chen, Xiao, Bihong Huang, and Dezhu Ye. 2018. The role of punctuation in P2P lending: Evidence from China. *Economic Modelling* 68:634-643.

Chen, Xiao-hong, Fu-jing Jin, Qun Zhang, and Li Yang. 2016. Are investors rational or perceptual in P2P lending? *Information Systems and e-Business Management* 14 (4):921-944.

Chenavaz, Régis. 2017. Dynamic quality policies with reference quality effects. *Applied Economics* 49 (32):3156-3162.

Choi, Hwan-Sik, and Ron A. Laschever. 2018. The credit card debt puzzle and noncognitive ability. *Review of Finance* 22 (6):2109-2137.

Cohn, Deborah Y., and Valerie L. Vaccaro. 2006. A study of neutralisation theory's application to global consumer ethics: P2P file-trading of musical intellectual property on the internet. *International Journal of Internet Marketing and Advertising* 3 (1):68-88.

Cole, Shawn A., Anna L. Paulson, and Gauri K. Shastry. 2016. High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of human resources* 51 (3):656-698.

Collins, J. M., and Elizabeth R. Odders-White. 2015. A framework for developing and testing financial capability education programs targeted to elementary schools. *The journal of economic education* 46 (1):105-120.

Cunha, M. R., Bart M. Lambrecht, and Grzegorz Pawlina. 2011. Household Liquidity and Incremental Financing Decisions: Theory and Evidence. *Journal of Business Finance & Accounting* 38 (7-8):1016-1052.

Datta, Shounak, Supritam Bhattacharjee, and Swagatam Das. 2018. Clustering with missing features: a penalized dissimilarity measure based approach. *Machine Learning* 107 (12):1987-2025.

Dawsey, Amanda E. 2015. State bankruptcy laws and the responsiveness of credit card demand. *Journal of Economics and Business* 81:54-76.

Diallo, Boubacar, and Omar Al-Titi. 2017. Local growth and access to credit: Theory and evidence. *Journal of Macroeconomics* 54:410-423.

Druedahl, Jeppe, and Casper N. Jørgensen. 2018. Precautionary borrowing and the credit card debt puzzle. *Quantitative Economics* 9 (2):785-823.

Duca, John V., and Stuart S. Rosenthal. 1993. Borrowing Constraints, Household Debt, and Racial Discrimination in Loan Markets. *Journal of Financial Intermediation* 3 (1):77-103.

Duygan-Bump, Burcu, Alexey Levkov, and Judit Montoriol-Garriga. 2015. Financing constraints and unemployment: Evidence from the Great Recession. *Journal of Monetary Economics* 75:89-105.

Eisfeldt, Andrea L., and Adriano A. Rampini. 2007. New or used?

Investment with credit constraints. *Journal of Monetary Economics* 54 (8):2656-2681.

Ennaceur, Amel, Zied Elouedi, and Eric Lefevre. 2016. Belief AHP Method - AHP Method with the Belief Function Framework. *International Journal of Information Technology and Decision Making* 15 (3):553-573.

Ezequiel L Opez-Rubio, Esteban J. Palomo Francisco. 2018. Unsupervised Learning by Cluster Quality Optimization. *Information Sciences*.

Fan, Bojun, Hannah Ji, June Wei, and Sherwood Lambert. 2018. Development of tactical solutions for the e-credit card issuing industry. *International Journal of Accounting & Information Management* 26 (1):115-131.

Fan, Lu, and Swarn Chatterjee. 2017. Borrowing Decision of Households: An Examination of the Information Search Process. *Journal of Financial Counseling and Planning* 28 (1):95-106.

Feng, Yan, Xinlu Fan, and Yeujun Yoon. 2015. Lenders and borrowers' strategies in online peer-to-peer lending market: An empirical analysis of ppdai.com. *Journal of Electronic Commerce Research* 16 (3):242-260.

Feng, Yi, Yinping Mu, Benyong Hu, and Arun Kumar. 2014. Commodity options purchasing and credit financing under capital constraint. *International Journal of Production Economics* 153:230-237.

Gan, Christopher E. C., David A. Cohen, Baiding Hu, Minh C. Tran, Weikang Dong, and Annie Wang. 2016. The relationship between credit card attributes and the demographic characteristics of card users in China. *International Journal of Bank Marketing* 34 (7):966-984.

García, María J. R. 2013. FINANCIAL EDUCATION AND BEHAVIORAL FINANCE: NEW INSIGHTS INTO THE ROLE OF INFORMATION IN FINANCIAL DECISIONS. *Journal of Economic Surveys* 27 (2):297-315.

Getachew, Yoseph Yilma. 2016. Credit constraints, growth and inequality dynamics. *Economic Modelling* 54:364-376.

Grant, C. 2007. Estimating credit constraints among US households. *Oxford Economic Papers* 59 (4):583-605.

Greiner, Martina E., and Hui Wang. 2010. Building Consumer-to-Consumer Trust in E-Finance Marketplaces: An

Empirical Analysis. *International Journal of Electronic Commerce* 15 (2):105-136.

Guo, Yanhong, Wenjun Zhou, Chunyu Luo, Chuanren Liu, and Hui Xiong. 2016. Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research* 249 (2):417-426.

Hai, Rong, and James J. Heckman. 2017. Inequality in human capital and endogenous credit constraints. *Review of Economic Dynamics* 25:4-36.

Hall, Robert E. 1978. Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy* 86 (6):971-987.

Han, Lu, Jing J. Xiao, and Zhi Su. 2018. Financing knowledge, risk attitude and P2P borrowing in China. *International Journal of Consumer Studies*.

HAN, SONG, and GENG LI. 2011. Household Borrowing after Personal Bankruptcy. *Journal of Money, Credit and Banking* 43 (2):491-517.

Havrylchyk, Olena, and Marianne Verdier. 2018. The Financial Intermediation Role of the P2P Lending Platforms. *Comparative Economic Studies* 60 (1):115-130.

Hayashi, Fumio. 1985. Tests for Liquidity Constraints: A Critical Survey. *National Bureau of Economic Research Working Paper Series* No. 1720.

Hedlund, Aaron. 2018. Credit constraints, house prices, and the impact of life cycle dynamics. *Economics Letters* 171:202-207.

Hu, Bao Q., and Heung Wong. 2013. Generalized interval-valued fuzzy rough sets based on interval-valued fuzzy logical operators. *International Journal of Fuzzy Systems* 15 (4):381-391.

Huang, Xiaohui, Yunming Ye, Huifeng Guo, Yi Cai, Haijun Zhang, and Yan Li. 2014. DSKmeans: A new kmeans-type approach to discriminative subspace clustering. *Knowledge-Based Systems* 70:293-300.

Huh, Jaeyung, Woojin Chang, Jaeyong Lee, and Junghoon Lee. 2010. Samsung card lending model. *European Journal of Operational Research* 207 (1):492-498.

Hung, Fu-Sheng. 2005. Credit rationing and capital accumulation with investment and consumption loans revisited. *Journal of Development Economics* 78 (2):322-347.

Huston, Sandra J. 2012. Financial literacy and the cost of borrowing. *International Journal of Consumer Studies* 36 (5):566-572.

Insler, Michael, James Compton, and Pamela Schmitt. 2016. The investment decisions of young adults under relaxed borrowing constraints. *Journal of Behavioral and Experimental Economics* 64:106-121.

Jaffee, Dwight M., and Thomas Russell. 1976. Imperfect Information, Uncertainty, and Credit Rationing*. *The Quarterly Journal of Economics* 90 (4):651-666.

Jiang, Cuiqing, Zhao Wang, Ruiya Wang, and Yong Ding. 2018. Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. *Annals of Operations Research* 266 (1-2):511-529.

Jiang, Cuixia, Qifa Xu, Weiming Zhang, Mengting Li, and Shanlin Yang. 2018. Does automatic bidding mechanism affect herding behavior? Evidence from online P2P lending in China. *Journal of Behavioral and Experimental Finance* 20:39-44.

Johnson, Susan, Arvind Ashta, and Djamchid Assadi. 2010. Online or offline? the rise of "peer-to-peer" lending in microfinance. *Journal of Electronic Commerce in Organizations* 8 (3):26-37.

JOIREMAN, JEFF, JEREMY KEES, and DAVID SPROTT. 2010. Concern with Immediate Consequences Magnifies the Impact of Compulsive Buying Tendencies on College Students' Credit Card Debt. *The Journal of Consumer Affairs* 44 (1):155-178.

Kaboski, Joseph P., Molly Lipscomb, and Virgiliu Midrigan. 2014. The Aggregate Impact of Household Saving and Borrowing Constraints: Designing a Field Experiment in Uganda. *The American Economic Review* 104 (5):171-176.

Kehoe, Timothy J., and David K. Levine. 1993. Debt-Constrained Asset Markets. *The Review of Economic Studies* 60 (4):865-888.

Keys, Benjamin J., and Jialan Wang. 2019. Minimum payments and debt paydown in consumer credit cards. *Journal of Financial Economics* 131 (3):528-548.

Khan, Hashmat, and Jean-François Rouillard. 2018. Household borrowing constraints and residential investment dynamics. *Journal of Economic Dynamics and Control* 95:1-18.

Kim, Jinhee, Michael S. Gutter, and Taylor Spangler. 2017. Review of family financial decision making: Suggestions for future research and implications for financial education. *Journal of Financial*

Counseling and Planning 28 (2):253-267.

Kim, Kyoung Tae, Melissa J. Wilmarth, and Shinae Choi. 2016. Credit Use of U.S. Households after the Great Recession: The Role of Credit Constraint. *Family and Consumer Sciences Research Journal* 44 (3):280-294.

Kroszner, Randall S., and Philip E. Strahan. 1999. What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions. *The Quarterly Journal of Economics* 114 (4):1437-1467.

Kumar, K. M., and ARM Reddy. 2017. An efficient k -means clustering filtering algorithm using density based initial cluster centers. *Information Sciences*: 286-301.

Levin, Irwin P., Elaine A. Bossard, Gary J. Gaeth, and Haoyang Yan. 2014. The combined role of task, child's age and individual differences in understanding decision processes. *Judgment and Decision Making* 9 (3):274-285.

Li, Changsheng, Liqiong Lin, and Christopher E. C. Gan. 2016. China credit constraints and rural households' consumption expenditure. *Finance Research Letters* 19:158-164.

Li, Zhihong, Lanteng Wu, and Hongting Tang. 2018. Optimizing the Borrowing Limit and Interest Rate in P2P System: From Borrowers' Perspective. *Scientific Programming* 2018:1-14.

Lin, Mingfeng, Nagpurnanand R. Prabhala, and Siva Viswanathan. 2013. Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending. *Management Science* 59 (1):17-35.

Liu, Chien-Liang, Tao-Hsing Chang, and Hsuan-Hsun Li. 2013. Clustering documents with labeled and unlabeled documents using fuzzy semi-Kmeans. *Fuzzy Sets and Systems* 221:48-64.

Liu, He, Han Qiao, Shouyang Wang, and Yuze Li. 2019. Platform Competition in Peer-to-Peer Lending Considering Risk Control Ability. *European Journal of Operational Research* 274 (1):280-290.

Lu, Meilian, Zhen Qin, Yiming Cao, Zhichao Liu, and Mengxing Wang. 2014. Scalable news recommendation using multi-dimensional similarity and Jaccard–Kmeans clustering. *Journal of Systems and Software* 95:242-251.

Lusardi, Annamaria, and Olivia S. Mitchell. 2014. The economic importance of financial literacy: theory and evidence: Theory and evidence. *Journal of economic literature* 52 (1):44.

Macharis, Cathy, Johan Springael, Klaas De Brucker, and Alain Verbeke. 2004. PROMETHEE and AHP: The design of operational synergies in multicriteria analysis.: Strengthening PROMETHEE with ideas of AHP. *European Journal of Operational Research* 153 (2):307-317.

Malkoc, Selin A., Gal Zauberman, and Canan Ulu. 2005. Consuming Now or Later? The Interactive Effect of Timing and Attribute Alignability. *Psychological Science* 16 (5):411-417.

Mikhed, Vyacheslav, and Michael Vogan. 2018. How data breaches affect consumer credit. *Journal of Banking & Finance* 88:192-207.

Mild, Andreas, Martin Waitz, and Jürgen Wöckl. 2015. How low can you go? — Overcoming the inability of lenders to set proper interest rates on unsecured peer-to-peer lending markets. *Journal of Business Research* 68 (6):1291-1305.

Nakamura, Leonard I., and Kasper Roszbach. 2017. Credit ratings, private information, and bank monitoring ability. *Journal of Financial Intermediation*.

Norvilitis, Jill M., Michelle M. Merwin, Timothy M. Osberg, Patricia V. Roehling, Paul Young, and Michele M. Kamas. 2006. Personality Factors, Money Attitudes, Financial Knowledge, and Credit-Card Debt in College Students. *Journal of Applied Social Psychology* 36 (6):1395-1413.

Parahoo, Sanjai K. 2012. Credit where it is due: drivers of loyalty to credit cards. *International Journal of Bank Marketing* 30 (1):4-19.

Paul, Manimoy, James Nolan, and Andrea Smith-Hunter. 2017. Financially at Risk: A look at College Students. Credit Card Usage. *Journal of Higher Education Theory and Practice* 17 (5):110-128.

Ricaldi, Laura, Michael S. Finke, and Sandra J. Huston. 2013. Financial literacy and shrouded credit card rewards. *Journal of Financial Services Marketing* 18 (3):177-187.

Ruiz Tagle, Jaime, and Francis Vella. 2016. Borrowing Constraints and Credit Demand in a Developing Economy. *Journal of Applied Econometrics* 31 (5):865-891.

Saaty, Thomas L. 2013. The Modern Science of Multicriteria Decision Making and Its Practical Applications: The AHP/ANP Approach. *Operations Research* 61 (5):1101-1118.

Saaty, Thomas L. 2003. Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal of Operational Research* 145 (1):85-91.

Sarkar, Sudipto. 2016. Consumer welfare and the strategic choice of price cap and leverage ratio. *Quarterly Review of Economics and Finance* 60:103-114.

Singh, Shweta, David H. Rylander, and Tina C. Mims. 2018. Understanding credit card payment behavior among college students. *Journal of Financial Services Marketing* 23 (1):38-49.

Song, Pingfan, Yunzhi Chen, Zhixiang Zhou, and Huaqing Wu. 2018. Performance analysis of Peer-to-Peer online lending platforms in China. *Sustainability (Switzerland)* 10 (9):2987.

Srikanthverma, Ch, and V. Ranaprathap. 2011. Cashless Society: A Study of Trends and Issues in Credit Card Industry in India. *Management Prudence* 2 (2):93.

Steffes, Erin M., B. P. S. Murthi, and Ram C. Rao. 2011. Why are some modes of acquisition more profitable? A study of the credit card industry. *Journal of Financial Services Marketing* 16 (2):90-100.

Steijvers, Tensie, Wim Voordeckers, and Koen Vanhoof. 2010. Collateral, relationship lending and family firms. *Small Business Economics* 34 (3):243-259.

Stern, Caroline, Mikko Makinen, and Zongxin Qian. 2017. FinTechs in China – with a special focus on peer to peer lending. *Journal of Chinese Economic and Foreign Trade Studies* 10 (3):215-228.

Stiglitz, Joseph E., and Andrew Weiss. 1981. Credit Rationing in Markets with Imperfect Information. *The American Economic Review* 71 (3):393-410.

Tavares-Gärtner, Miguel, Paulo J. Pereira, and Elísio Brandão. 2018. Optimal contingent payment mechanisms and entrepreneurial financing decisions. *European Journal of Operational Research* 270 (3):1182-1194.

TELYUKOVA, IRINA A. 2013. Household Need for Liquidity and the Credit Card Debt Puzzle. *The Review of Economic Studies* 80 (3 (284)):1148-1177.

Ting, Kai Ming, Ye Zhu, Mark Carman, Yue Zhu, Takashi Washio, and Zhi-Hua Zhou. 2018. Lowest probability mass neighbour algorithms: relaxing the metric constraint in distance-based neighbourhood algorithms. *Machine Learning*.

Venkatesh, Viswanath, Michael G. Morris, and Phillip L. Ackerman. 2000. A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes. *Organizational Behavior and Human Decision Processes* 83 (1):33-60.

Verplanken, Bas, and Rob W. Holland. 2002. Motivated Decision Making: Effects of Activation and Self-Centrality of Values on Choices and Behavior. *Journal of Personality and Social Psychology* 82 (3):434-447.

Vigil, Virginija K. 2009. Models of an individual decision-making process related to ethical issues in business: The risk of framing effects. *International Journal of Business Governance and Ethics* 4 (3):264-279.

Wang, Junlong. 2017. Research on Financing and Decision from Micro Enterprises in China —Based on Internet Financial Model Perspective. *Open Journal of Business and Management* 5 (2):372-387.

Wang, Lili, Wei Lu, and Naresh K. Malhotra. 2011. Demographics, attitude, personality and credit card features correlate with credit card debt: A view from China. *Journal of Economic Psychology* 32 (1):179-193.

WILCOX, KEITH, LAUREN G. BLOCK, and ERIC M. EISENSTEIN. 2011. Leave Home Without It? The Effects of Credit Card Debt and Available Credit on Spending. *Journal of Marketing Research* 48:S78-S90.

Wong, King-Yin, and Michael Lynn. 2017. The easy-money effect: credit card spending and hard-work reminders. *Journal of Consumer Marketing* 34 (7):541-551.

Wu, Zhihao, Youfang Lin, Yiji Zhao, and Hongyan Yan. 2018. Improving local clustering based top-L link prediction methods via asymmetric link clustering information. *Physica A: Statistical Mechanics and its Applications* 492:1859-1874.

Zahra, Sobia, Mustansar Ali Ghazanfar, Asra Khalid, Muhammad Awais Azam, Usman Naeem, and Adam Prugel-Bennett. 2015. Novel centroid selection approaches for KMeans-clustering based recommender systems. *Information Sciences* 320:156-189.

Zhang, Hefu, Hongke Zhao, Qi Liu, Tong Xu, Enhong Chen, and Xunpeng Huang. 2018. Finding potential lenders in P2P lending: A Hybrid Random Walk Approach. *Information Sciences*

432:376-391.

Zhang, Nanfei. 2017. Rationality of Investors in P2P Online Lending Platform with Guarantee Mechanism: Evidence in China. *Journal of Applied Finance and Banking* 7 (3):121.

Zhao, Hongke, Yong Ge, Qi Liu, Guifeng Wang, Enhong Chen, and Hefu Zhang. 2017. P2P Lending Survey: Platforms, Recent Advances and Prospects. *ACM Transactions on Intelligent Systems and Technology (TIST)* 8 (6):1-28.

Zhao, Yi, Ying Zhao, and Inseong Song. 2009. Predicting New Customers' Risk Type in the Credit Card Market. *Journal of Marketing Research* 46 (4):506-517.

THE SURVEY OF HOUSEHOLD FINANCIAL SITUATIONS

1. Basic information

<table>
<tr><td rowspan="2">1A. gender</td><td>1</td><td colspan="3">male</td></tr>
<tr><td>2</td><td colspan="3">female</td></tr>
<tr><td rowspan="3">1B. marital status</td><td>1</td><td colspan="3">unmarried</td></tr>
<tr><td>2</td><td colspan="3">married</td></tr>
<tr><td>3</td><td colspan="3">other（such as divorced or widowed）</td></tr>
<tr><td rowspan="2">1C.age</td><td colspan="2">your age</td><td colspan="2"></td></tr>
<tr><td colspan="2">your spouse’s age</td><td colspan="2"></td></tr>
<tr><td rowspan="3">1D. working years (years)</td><td colspan="2"></td><td>you(N)</td><td>your spouse(P)</td></tr>
<tr><td colspan="2">years since getting your current job</td><td></td><td></td></tr>
<tr><td colspan="2">years since getting your current job</td><td></td><td></td></tr>
<tr><td rowspan="8">1E. you and your spouse’s occupations (single choice)</td><td>you (N)</td><td>your spouse (P)</td><td colspan="2"></td></tr>
<tr><td>1</td><td>1</td><td colspan="2">government agencies, party chiefs or senior officials</td></tr>
<tr><td>2</td><td>2</td><td colspan="2">management of enterprises or institutions</td></tr>
<tr><td>3</td><td>3</td><td colspan="2">government or corporate employees</td></tr>
<tr><td>4</td><td>4</td><td colspan="2">professional technicians or other professionals</td></tr>
<tr><td>5</td><td>5</td><td colspan="2">skilled worker</td></tr>
<tr><td>6</td><td>6</td><td colspan="2">self-employed</td></tr>
<tr><td>7</td><td>7</td><td colspan="2">freelance (refers to freelance writers, programmers, and other freelancers)</td></tr>
</table>

	8	8	other occupations (please specify:____________)
	9	9	unemployed
	10	10	retired
1F. you and your spouse's industries (single choice)	you (N)	your spouse (P)	
	1	1	manufacturing
	2	2	education and research, health care and social welfare
	3	3	government and social organizations
	4	4	architecture
	5	5	transport, storage and postal telecommunications
	6	6	extractive industries
	7	7	wholesale, retail, trading, and catering industries
	8	8	financial industry
	9	9	electricity, gas and water production
	10	10	real estate industry
	11	11	other (including leasing, geological survey, information service, etc.) (please specify: ____________)
1G. you and your spouse's education (single choice)	you (N)	your spouse (P)	
	1	1	junior high school and below
	2	2	high school and vocational school
	3	3	three-year or four-year college
	4	4	master's degree
	5	5	doctoral degree

1H. How many elderly people do you need to support?				
1I. Your child's age?（From young to old. If you do not have children, you can skip this question）	A	B	C	D

2. Family investment

<table>
<tr><td rowspan="2">2A. Do you have relatives engaged in the investment industry (stocks, bonds, funds, etc.)?</td><td colspan="2">1</td><td>yes</td></tr>
<tr><td colspan="2">2</td><td>no</td></tr>
<tr><td rowspan="7">2B. Does your family have investment experience with the following products?</td><td>yes</td><td>no</td><td></td></tr>
<tr><td>1</td><td>2</td><td>stock (A)</td></tr>
<tr><td>1</td><td>2</td><td>fund (B)</td></tr>
<tr><td>1</td><td>2</td><td>bonds (C)</td></tr>
<tr><td>1</td><td>2</td><td>foreign exchange (D)</td></tr>
<tr><td>1</td><td>2</td><td>value-added preserving products (gold, artwork, etc.) (E)</td></tr>
<tr><td>1</td><td>2</td><td>real estate (referring to investment property) (F)</td></tr>
<tr><td rowspan="2">2C. Read the question and give your answer</td><td>You(N)</td><td>Your spouse (P)</td><td rowspan="2">Suppose you throw a coin, and if its face is up, you get $2,000; otherwise, you get nothing if it's tails-up. If you sell this profit-making opportunity to others, how much would you ask for it?</td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="2">2D. Read the question and give your answer</td><td>You(N)</td><td>Your spouse (P)</td><td rowspan="2">Suppose you are forced to throw a coin once, and if its face is up you will not lose anything. If it's tails-up, you will lose $20,000. If you have the opportunity to pass this burden to others, how much would you be willing to pay?</td></tr>
<tr><td></td><td></td></tr>
</table>

2E. What levels of risk are you willing to take when you and your spouse invest?	You(N)	Your spouse (P)	
	5	5	take high risk for high returns
	4	4	take somewhat high risk for somewhat high returns
	3	3	take the average risk for the average return
	2	2	take low risk for low return
	1	1	unwilling to take any risk

2F. Given a 10-point risk mark for stocks and a 0-point risk mark for bank savings, what risk level do you think the following financial products should be rated at?		fund	A
		bond	B
		foreign exchange	C
		value-preservingproducts (gold, artwork, etc.)	D
		real estate (referring to investment property)	E

2G. Have you or your spouse completed high school education in economics or management?		yes	Whether you or your partner have a higher education degree in finance		yes	1
					no	2
		no				3

2H. How much in available funds did you have for investing last year? （） yuan

2I. How much income did your family have last year? （） yuan

3. Family financing

3A. If you need to borrow money, who is your "preferred" loaner? **(single choice)**	1	relatives and/or friends
	2	bank
	3	non-bank formal financial institutions
	4	private lending institutions and individuals
3B. For your family, how would it be to raise 100,000 yuan and the term is 1 year? **(single choice)**	1	very difficult
	2	difficult
	3	so-so
	4	easy
	5	very easy
3C. If your family borrows money from relatives and friends, what will be the interest rate you pay to them? **(single choice)**	1	no interest
	2	lower than the deposit interest rate
	3	deposit interest rates
	4	between deposit interest rate and loan interest rate
	5	loan interest rate
	6	higher than loan interest rate
3D. If you borrow money from friends and relatives,	1	mortgage
	2	guarantees

what arrangement will be mainly used? **(single choice)**	3	IOU (I owe you)
	4	oral agreement
	5	other (please specify: ____________)
3E. If you borrow money from friends and relatives, what is the longest term? **(single choice)**	1	within a year
	2	one or two years
	3	three years to four years
	4	more than four years
	5	no specific term
	6	Do not borrow money from friends and relatives

3F. If you borrow money from a bank or financial institution, what is the maximum interest rate you can tolerate?	4%	6%	8%	10%	15%	20%
	1	2	3	4	5	6

3G. Do you know the following loan products from the commercial banks?	Not at all	A little	Gene	A lot	Very much
Housing loan**(A)**	1	2	3	4	5
Car loan **(B)**	1	2	3	4	5
Decoration loan**(C)**	1	2	3	4	5
Education loan**(D)**	1	2	3	4	5
Business loan **(E)**	1	2	3	4	5
Large commodity loan**(F)**	1	2	3	4	5
3H. In your opinion, how difficult is it to ask your bank to apply for the following loans?	very difficult	difficult	so-so	easy	very easy
Housing loan**(A)**	1	2	3	4	5
Car loan **(B)**	1	2	3	4	5
Decoration loan**(C)**	1	2	3	4	5
Education loan**(D)**	1	2	3	4	5
Business loan **(E)**	1	2	3	4	5
Large commodity loan**(F)**	1	2	3	4	5

<table>
<tr><td rowspan="2">3I. How much debt can you afford equivalent to times of your annual household income?</td><td><1</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>>10</td></tr>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr>
<tr><td rowspan="5">3J. If you need a loan, what kind of property will your family use as collateral?</td><td colspan="2">A</td><td colspan="9">deposit slip</td></tr>
<tr><td colspan="2">B</td><td colspan="9">stocks, bonds, funds and other investment financial assets</td></tr>
<tr><td colspan="2">C</td><td colspan="9">insurance</td></tr>
<tr><td colspan="2">D</td><td colspan="9">real estate</td></tr>
<tr><td colspan="2">E</td><td colspan="9">Other (please specify: ____________)</td></tr>
<tr><td rowspan="3">3K. Have you ever borrowed money via the internet? (single choice)</td><td colspan="2">1</td><td colspan="9">applied for it and got it (please answer question 3M)</td></tr>
<tr><td colspan="2">2</td><td colspan="9">applied for it but did not get it</td></tr>
<tr><td colspan="2">3</td><td colspan="9">never applied (please answer question 3L)</td></tr>
</table>

3L. What concerns you the most about borrowing via the internet?	1	high interest
	2	more collateral
	3	not enough time to repay
	4	personal privacy
	5	other(please specify: ____________)
3M. Do you shop around when selecting an internet financing platform?	1	yes (please answer question 3N)
	2	no
3N. What is the most important indicator when you select a financing platform for borrowing? (single choice)	1	relatively low interest
	2	short time to wait for money
	3	satisfying the total quota
	4	convenient to apply
	5	available for extension
	6	no need for collateral
	7	other(please specify: ____________)

3O. Do you know that P2P always has a financing fee? (Financing fee is charged one time when the deal is fixed.)	1	yes			
	2	no			
3P. Do you know that P2P always has a .management fee? **(**Management fee is charged every month)	1	yes			
	2	no			
3Q. Please rank these indicators for borrowing. (1 refers to not important at all; 9 refers to most important)	Convenience ( )	Cost ( )	Quota ( )	Speed ( )	Repayment ( )

3R. Rank Private vs. Bank on the same indicators as 3Q. (Private means borrowing from relatives and/or friends; Bank means borrowing from banks.)					
3S. Rank Private vs. Nonbank on the same indicators as 3Q. (Private means borrowing from relatives and/or friends; Nonbank means borrowing from financial institutions, including internet financing.)					

3T. Rank Private vs. Others on the same indicators as 3Q. (Private means borrowing from relatives and/or friends; Others mean borrowing from channels which are not listed above.)					
3U. Rank Bank vs. Nonbank on the same indicators as 3Q.					
3V. Rank Bank vs. Others on the same indicators as 3Q.					
3W. Rank Nonbank vs. others on the same indicators as 3Q.					