

# **Occupational Identity Discrimination in Peer-to-Peer Lending**

**Xiangbo Cui**<sup>1</sup>

## **Abstract**

Using data collected from a leading P2P platform in China, this paper empirically tests the discrimination of investors on the occupational identity of borrowers in online lending. I find that P2P investors discriminate against borrowers who are salary earners in terms of occupational identity while preferring borrowers who are private entrepreneurs. Moreover, this kind of discrimination can be found in borrowers both with high credit ratings and low credit ratings. The findings also imply that the occupational identity of borrowers plays a role of moderating the relationship between the credit rating and the probability of successful funding. Compared with private entrepreneurs, credit rating has a weaker effect on the probability of successful funding among salary earners. However, after examining the default rate, the results show that the default rate of private entrepreneurs is significantly higher than that of salary earners. This indicates that the discrimination of P2P investors is not based on some rational economic roots and is a kind of inefficient tasted-based discrimination.

**JEL classification numbers:** H 74, J 71

**Keywords:** occupational identity, peer-to-peer lending, probability of successful funding, borrowing cost, discrimination

---

<sup>1</sup> Tsinghua University, PBC School of Finance, 43 Chengfu Road, Beijing 100083, China

## **1. Introduction**

Social discrimination refers to the unfair or unequal treatment of some people in light of their membership of a particular group. There are many causes of social discrimination, but the fundamental cause is to offer unfair and exclusive differentiated treatment to certain vulnerable groups out of identity classification rather than quality, such as the discrimination against women and patients. According to the theory of social identity, individuals tend to construct their own and other people's identities through social category, which means that people always consciously classify and evaluate social groups. In this process, people distinguish between members of their own community and those of other communities, and give themselves a community-based social identity. Such classification is ubiquitous, and different social identities are adopted for different situations. According to Tajfel et al (1973), these basic classifications are sufficient to create discrimination. Among multiple social identities, professional identity has become an important label of individuals in social life. Discrimination based on professional identity has always existed, and has become an important research field of sociology and economics. This paper hopes to continue to carry out research in this field. Specifically, this paper mainly focuses on investors' discrimination against borrowers based on occupational identity in online lending.

Credit discrimination is a pervasive phenomenon in the credit market. Some borrowers are always inferior to others in the process of credit application. The rejection rate of credit application for these people is far higher than that of other borrowers. Credit discrimination is divided into credit discrimination against enterprises and that against individuals. At the corporate level, owing to the different risk profiles of different types of firms and the credit rationing of banks, credit discrimination against firms by banks has always existed, especially in transition economies. For example, state-owned enterprises have easier access to bank loans than private enterprises (Brandt and Li, 2003); politically-connected private companies are also more likely to obtain loans (Claessens, Feijen, & Laeven, 2008; Li et al., 2008). At the individual level, credit discrimination is even more the focus of social attention, because credit availability is closely related to household welfare. In previous studies, a lot of literature has explored credit discrimination based on gender (Bellucci, Borisov, & Zazzaro, 2010; Alesina, Lotti, & Mistrulli, 2013; Cozarenco, & Szafarz, 2018), race (Cavalluzzo, & Cavalluzzo, 1998; Blanchflower, Levine, & Zimmerman, 2003; Blanchard, Zhao, & Yinger, 2008), and region of borrowers (Lin, & Viswanathan, 2016). In fact, aside from these pieces of information, borrowers also have another important type of information mastered by the lender, namely occupational identity, such as wage earners, the unemployed, students, managers, and private entrepreneurs. Both in traditional bank lending, and in new types of lending such as P2P and payday loan, the occupation of borrowers is always an important basis for lending decisions. The reason is that occupational identity usually represents the stability of a profession, and to a certain extent reflects the level of individual wealth. These are critical

factors influencing borrowers' ability to repay loans. In academic research, due to the lack of data on microcredit, there are few studies on whether credit discrimination based on occupational identity exists or not. This paper has obtained over 300,000 pieces of loan data from a leading P2P platform in China, providing an opportunity to study the issue. Therefore, this paper intends to study whether credit discrimination based on the occupational identity of borrowers exists or not in online lending.

P2P lending appeared first in the UK in 2005. Different from traditional bank credit, P2P lending allows borrowers to directly obtain capital from individual investors and is a financial service model that directly realizes lending behaviors on the Internet without financial institutions. Potential investors on the platform decide whether or not to make the investment and the amount of investment based on the information of each loan listing. If the total amount of desired investment of all investors for a loan has reached the predetermined borrowing amount in a prescribed period, the borrowing is successful. Or else, the borrowing fails. P2P online lending provides new solutions to easing the financing difficulties of individuals and small and micro businesses. Meanwhile, it has also provided wider investment channels for social idle funds, so upon its birth, it has won greater attention and enjoyed fast development.

However, not everyone can successfully obtain loans from P2P platforms. In fact, the probability of successful funding on P2P platforms is not high. For example, as for Prosper, the largest P2P platform in the world, the probability of successful funding is only around 10%. An important reason for such a phenomenon is investors' lack of trust on borrowers. Since investors cannot make face-to-face contact with borrowers or actively collect relevant information, serious information asymmetry exists between borrowers and investors. Investors can only make decisions based on the information disclosed by borrowers, including both soft information and hard information (Herzenstein, Sonenshein, & Dholakia, 2011; Lin et al, 2013; Dorfleitner et al, 2016; Chen, Zhou, & Wan, 2016). Hard information differs from soft information mainly in terms of quantifiability or subjectivity and objectivity. Hard information usually refers to information that can be easily converted into a digital format and does not require subjective judgment, such as age, gender, financial condition, and credit score, etc. Soft information mainly refers to information that is difficult to be converted into digital format or requires subjective judgment in the process of identification, such as language description, photos and images, and social identity, etc. Moreover, in soft information, occupational identity is the information directly reflecting the income of an individual, wealth status, and social stratum of individuals. Through occupational information, we can more directly understand the loan repayment ability of a borrower. So, will investors regard occupational identity as a key signal to screen good borrowers? Will investors discriminate against borrowers of some occupational identities? This is the focus of the paper.

This paper first studies whether there is discrimination against the occupational identity of borrowers in P2P lending. Among the borrowers of P2P platforms, there

are two most typical occupational identities: the wage earner and the private entrepreneur. The former refers to the people who rely on working to earn salary income, while the later refers to investors and owners of private enterprises. In traditional bank credit, private entrepreneurs, especially owners of small enterprises, are less likely to get bank loans than salaried workers with stable occupations. One of the important reasons is that banks are worried about the repayment ability of business owners. Because of the uncertainties in the enterprise operation, they may not be able to repay bank loans when facing higher operational risks. The wage earners have more stable cash flow than private entrepreneurs because of their stable wage income. However, in the new-type P2P lending, whether these two identities of borrowers will be prejudiced by investors has not been fully studied. If borrowers with a certain occupational identity have a higher probability of successful funding, it means that the investors have a higher willingness to invest in the borrower with this occupational identity. Therefore, by examining the success rate of borrowers with different occupational identities, the purpose of testing whether investors have credit discrimination against occupational identities can be achieved. This paper finds that private entrepreneurs are more likely to get loans than the wage earners, which is quite different from that in traditional bank credit. It indicates that investors on the P2P platform prefer private entrepreneurs and discriminate against salary earners.

In order to further test the phenomenon that investors prefer private entrepreneurs while discriminating against the salary earners, this paper makes an in-depth analysis from three perspectives. Firstly, I examine whether there is discrimination against the salary earners among borrowers with different credit ratings. Since most of the P2P borrowers are difficult to obtain loans from traditional credit channels, this paper is concerned that the difference in credit ratings may interfere with previous conclusions. However, the results show that for borrowers with the HR credit rating, investors also prefer private entrepreneurs. For borrowers with higher credit ratings than HR, the same results are found, which shows that discrimination of investors does not vary according to the difference of borrower's credit rating. Secondly, the interest rate levels of the salary earner and the private entrepreneur are also tested. The way to determine the lending rate on Renrendai.com is that first, the platform determines the credit rating of borrowers based on their application information, and then provides them with an interest rate interval corresponding to the credit rating, in which borrowers choose the interest rate independently according to their borrowing demand. Therefore, by examining the interest rate of each loan listing, on the one hand, whether the difference of interest rate level causes discrimination can be examined, on the other hand, whether the overall credit risk of the salary earner and the private entrepreneur is different by can also be judged. Empirical results show that private entrepreneurs obtain lower interest rates. Finally, this paper examines the moderating effect of occupational identity on the relationship between credit rating and probability of successful funding. If investors do discriminate against the borrower's occupational identity, such discrimination should affect the role of other variables that have an impact on the probability of

successful funding. Among other variables affecting the probability of successful funding, the borrower's credit rating is the most important. This paper finds that the identity of private entrepreneur will enhance the impact of credit rating on the probability of successful funding.

After examining the investor's discrimination against the borrower's occupational identity, I further explore whether the investor's discrimination is rational or irrational. In economic theories, scholars distinguish two kinds of discrimination: one is the efficient statistical discrimination proposed by Phelps (1972) and Arrow (1973), the other is the inefficient tasted-based discrimination proposed by Becker (1957). Under the context of P2P lending, the former refers to the fact that some groups have inherent economic roots in their easier access to borrowing than others, such as higher rates of return or lower default rates offered by them; the later refers to the fact that investors have no reasonable economic roots for credit discrimination against certain groups but based entirely on the likes and dislikes of investors. This paper examines the default rate of private entrepreneurs and the salary earners to determine whether the discrimination of investors is the efficient statistical discrimination or the inefficient tasted-based discrimination. The results show that the default rate of private entrepreneurs is significantly higher than that of the salary earners, that is, private entrepreneurs are more likely to default after successful borrowing, which indicates that although investors prefer private entrepreneurs to the salary earners, the default rate of the salary earners is not higher than that of private entrepreneurs. That is to say, investors' discrimination against the salary earners is irrational, being the inefficient tasted-based discrimination. These findings provide new empirical evidence for understanding the investment behaviors of Chinese P2P investors and the credit discrimination in P2P lending.

The structure of this paper is as follows: the first part is the introduction; the second part elaborates the theoretical mechanism and puts forward the research hypotheses; the third part presents the data sources; the fourth part puts forward the empirical method and models; the fifth part shows the empirical results; the sixth part talks about the robustness test; the seventh part discusses the conclusion and the limitations of this paper.

## **2. Hypothesis Development**

The first question concerned in this paper is whether P2P investors will discriminate against the occupational identity of borrowers. The occupational identity of borrowers in the loan data of this paper can be divided into two categories of the wage earner and the private entrepreneur, so more precisely, this paper is concerned about whether investors will discriminate against borrowers with the occupational identity of the wage earner or the private entrepreneur. In the context of P2P lending, there are two types of situations in which discrimination occurs. One is the stereotype of investors based on past information or experience. Generally speaking, private entrepreneurs face the business operation risks, and the stability of cash flow

would be worse than that of the wage earner. Moreover, bankruptcy or default incidents of some enterprises will also deepen people's inherent impression that private entrepreneurs are facing higher risks. Relatively speaking, the wage earners have a stable salary, which ensures the stability of their cash flow. This stereotype may lead to prejudice of investors against private entrepreneurs. Becker (1957) believes that prejudice is the root of discrimination. As a result, the prejudice against private entrepreneurs brought about by past experience or recognition of investors may eventually lead to the preference for the wage earner and the discrimination against private entrepreneurs.

The second is based on the credit status and default risk of P2P borrowers. Most of the P2P borrowers are those who, with relatively low credit ratings and high default risk, can't get loans from traditional credit channels. Specifically, there may be differences in the credit risk between the wage earners and private entrepreneurs. Private entrepreneurs may find it more difficult to obtain loans through traditional banking channels than the wage earners. The traditional bank credit scoring system is more disadvantageous for private entrepreneurs with small scale and high operational risk to obtain loans, which is a question of universal concern around the world. In fact, the ability and willingness of private entrepreneurs to repay the loans may not be less, but they are limited by the existing bank rating system, which makes it impossible to obtain loans from banks. As a result, these private entrepreneurs seek loans through the P2P platforms. By contrast, the wage earners themselves are more likely to get loans through traditional credit channels than private entrepreneurs because of their stable salaries; the ones who borrow money from the P2P platforms are likely to be borrowers with low credit level, insufficient repayment ability and even less repayment willingness. That is to say, the wage earner borrowers may have the problem of adverse selection, while the overall credit level and the possibility of repayment on time of private entrepreneurs on the P2P platform may be higher than that of the wage earners. As a result, P2P investors may also prefer private entrepreneurs to the wage earners. Compared with the first situation, the second is closer to the fact of P2P lending, that is, private entrepreneurs may have easier access to loans.

H1: Borrowers with occupational identity as the private entrepreneur are more likely to get loans successfully than those as the wage earner.

On the basis of the above hypothesis, this paper continues to analyze this discrimination from three perspectives: the borrower's credit rating, the interest rate and the moderating effect of occupational identity. If investors do discriminate against the borrower's occupational identity, the discrimination may exist in borrowers of different credit ratings, or only in borrowers with specific credit ratings. Renrendai.com divides the borrower's credit rating into seven levels: AA、A、B、C、D、E and HR. Most borrowers' credit rating is HR, which is also consistent with the fact that most borrowers of P2P platforms are difficult to obtain loans from

traditional credit channels. Because the risk level of HR borrowers is quite different from that of other credit-rating borrowers, most previous literature divide borrowers into HR and non-HR groups according to their credit ratings. Consistent with previous literature, this paper examines whether discrimination based on occupational identity exists in both HR borrowers and non-HR borrowers.

H2: Borrowers with occupational identity as the private entrepreneur are more likely to get loans successfully than those as wage earner, no matter they are HR borrowers or non-HR borrowers.

The interest rate in loan listings comprehensively reflects the risk level of borrowers. Interest rate is determined through providing a narrow interest rate range for the applicant by Renrendai.com according to their credit ratings, and then the applicant chooses the interest rate he/she wants to pay within this range. Therefore, the interest rate can also be used to judge the difference in borrowing costs that the salary earners and the private entrepreneurs are willing to pay, as well as the risk status of these two groups of borrowers. If the overall credit risk of private entrepreneurs on the P2P platform is lower than that of the salary earners, the interest rate offered by private entrepreneurs may also be lower than that by the salary earner.

H3: The interest rate of loan listings from private entrepreneurs is lower than that from the salary earners.

Occupational identity belongs to the soft information in loan application. In addition to the occupational identity, there are other soft and hard information, such as borrowing purpose, descriptive information of each loan listing, gender, age and credit rating, that can affect the probability of successful funding, among which the credit rating given by the P2P lending platform to borrowers is the most important. A large number of studies have shown that borrowers with higher credit ratings are more likely to succeed in obtaining loans. Since investors discriminate against the occupational identity of borrowers, will it affect the relationship between other information and the probability of successful funding? Namely, does this discrimination have the moderating effect on the impact of other information on the probability of successful funding? This paper focuses on how discrimination affects the impact of credit rating on the probability of successful funding. It can be reasonably inferred that if investors prefer private entrepreneurs and discriminate against the salary earners, the impact of credit rating on the probability of successful funding is certainly different for these two types of borrowers. For private entrepreneurs, the impact of credit rating on the probability of successful funding may be greater, that is, the improvement of the probability of successful funding will be greater if the credit rating is increased by one unit.

H4: For borrowers with occupational identity as the private entrepreneur, credit rating has a stronger impact on the probability of successful funding.

The last question concerned in this paper is that if investors do have credit discrimination based on the occupational identity of borrowers, is this discrimination efficient statistical discrimination or inefficient tasted-based discrimination? If there are reasonable economic reasons for the discrimination of investors, that is, private entrepreneurs have lower default rate or the salary earners have the higher, such discrimination is rational and efficient statistical discrimination. If the reasons for investors' discrimination do not have reasonable economic roots, but only caused by investor's prejudice against private entrepreneurs or the salary earners, and in fact, there is no difference between these two groups or the default rate of private entrepreneurs is higher, then this kind of discrimination is ineffective preference discrimination. According to the previous analysis, P2P borrowers generally have low credit ratings. However, compared with the salary earners, the credit status and repayment ability of private entrepreneurs may be better. So, the default rate may be lower among private entrepreneurs. Namely, if investors prefer private entrepreneurs and discriminate against the salary earners, this discrimination is more likely to be the efficient statistical discrimination.

H5: The default rate of borrowers with occupational identity as the private entrepreneur is lower than that of borrowers with occupational identity as the salary earner.

### **3. Data and Variable Definitions**

#### **3.1 Data**

I use loan data from the P2P lending platform "renrendai.com". Established in 2010, it is one of the most influential P2P lending platforms in China. At present, most of the literature on P2P lending based on Chinese data is using the loan listings of "renrendai.com", the earliest date of which can be traced back to October 2010. In order to avoid the interference of some test loan applications in the early stage of platform development, I use loan listings between January 1, 2011 and December 31, 2014 in this paper. Within the time range, 314646 consummated and failed loan listings are finally obtained.

#### **3.2 Variable Definitions and Summary Statistics**

Table 1 shows variable definitions, and Table 2 provides summary statistics of borrowers' attributes and loan characteristics.

**Table 1: Variable Definitions**

Variable	Symbol	Description
Occupational identity	Identity	If the borrower is a private entrepreneur, the value of id is equal to 1; if the borrower is a wage earner, the value of id is equal to 0.
HR level	Risk Level	The borrower is divided into seven grades from AA to HR at "renrendai.com". If the borrower is at the HR level, the value is equal to 1, otherwise it is 0.
Address verification	Address_cert	If the borrower has passed the home address verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Social network verification	Network_cert	If the borrower has passed the social network verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Telephone verification	Mobile_cert	If the borrower has passed the telephone verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Identity card verification	Idcard_cert	If the borrower has passed the identity card verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Marriage verification	Marriage_cert	If the borrower has passed the marriage verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Work verification	Work_cert	If the borrower has passed the work verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Credit report verification	Credit_cert	If the borrower has passed the credit report verification of the "renrendai.com", the value is equal to 1, otherwise it is 0.
Delinquency	Default	If the loan is confirmed as bad debt by the platform, the value is equal to 1, otherwise it is 0.
Success	Success	If the loan is fully funded, the value is equal to 1, otherwise it is 0.
Male	Male	If the borrower is male, the value is equal to 1, otherwise it is 0.
Interest Rate	Interest	Interest rate of each loan;
Loan Amount	Amount	The amount requested of each loan
Loan Term	Term	The term of each loan

**Table 2: Summary Statistics**

Variable	Obs	Mean	Std	Min	Max
<i>Panel A: Borrower Attributes</i>					
Identity	314,646	0.260	0.438	0	1
Male	314,646	0.865	0.342	0	1
Risk Level	314,646	1.141	0.693	1	7
Mobile_cert	314,646	0.873	0.333	0	1
Idcard_cert	314,646	0.908	0.289	0	1
Address_cert	314,646	0.0623	0.242	0	1
Network_cert	314,646	0.00673	0.0817	0	1
Marriage_cert	314,646	0.450	0.497	0	1
Work_cert	314,646	0.759	0.428	0	1
Credit_cert	314,646	0.843	0.364	0	1
<i>Panel B: Loan Characteristics</i>					
Success	314,646	0.0649	0.246	0	1
Default	20,412	0.118	0.322	0	1
Interest Rate	314,646	15.08	3.569	3	24.40
Amount	314,646	68,080	111,198	3,000	1,000,000
Term	314,646	14.84	9.810	3	36

#### 4. Research Design

Firstly, this paper studies the impact of borrower's occupational identity on the probability of successful funding, that is, whether the investor discriminates against the borrower with an occupational identity of the wage earner or the private entrepreneur. In view of the data characteristics of this paper, I choose probit model, as shown in model (1).

$$\Pr(\text{success}_i = 1) = \Phi(\beta * \text{identity}_i + \text{Controls} + \varepsilon_i) \quad (1)$$

In this model,  $\text{success}_i$  is a dummy variable. If the loan is fully funded, the value is equal to 1, otherwise it is equal to 0. This paper mainly focuses on the coefficient  $\beta$  of  $\text{identity}_i$  which is a dummy variable representing the occupational identity of the borrower of the 1st loan. If the borrower is a private entrepreneur, the value of  $\text{identity}_i$  is equal to 1; if the borrower is a wage earner, the value of  $\text{identity}_i$  is equal to 0. If  $\beta$  is significantly positive, it indicates that private entrepreneurs are

more likely to get loans, that is, investors prefer private entrepreneurs, and discriminate against borrowers with the occupational identity as the wage earners. Other variables in model (1) include control variables and year fixed effects. The control variables include variables on borrower attributes and variables on loan characteristics. The variables on borrower characteristics include the borrower's gender, age and credit rating. The variables on loan characteristics cover the interest rate, loan amount, loan duration and some important information verifications. Among them, the information verifications refer to telephone verification, identity card authentication、marriage verification、credit report verification、work verification、address verification、social network verification, all of which are dummy variables. If they have been successfully authenticated, the value is equal to 1, otherwise it is equal to 0. Because the loan data in this paper cover the period from 2011 to 2014, I also include year fixed effect to control for factors changing each year.

In order to test the moderating effect of occupational identity on the relationship between credit rating and the probability of successful funding, this paper also adopts the probit model, as shown in model (2).

$$\Pr(\text{success}_i = 1) = \phi(\beta_0 * \text{risk level}_i + \beta_1 * \text{risk level}_i * \text{identity}_i + \varepsilon_i) \quad (2)$$

Risk level<sub>*i*</sub> represents the borrower's credit rating, with values ranging from 1 to 7, among which 1 representing the lowest while 7 representing the highest. If the coefficient of risk level<sub>*i*</sub> is significantly greater than 0, it can be explained that credit rating has an impact on the probability of successful funding. This paper pays attention to the coefficient  $\beta_1$  of the cross term  $\text{risk level}_i * \text{identity}_i$ . If the coefficient is greater than 0, it indicates that the occupational identity of the private entrepreneur can enhance the impact of credit rating on the probability of successful funding, that is, the relationship between credit rating and probability of successful funding will be affected by the occupational identity of borrowers.

In order to test the difference of borrowing costs that borrowers with different occupational identities are willing to pay, OLS regression is adopted as shown in (3).

$$\text{interest}_i = \beta * \text{identity}_i + \text{Controls} + \varepsilon_i \quad (3)$$

The explanatory variable interest<sub>*i*</sub> is the interest rate level of the 1st order, which is used to measure the borrowing cost of borrowers. The main concern of this paper is also the coefficient of  $\text{identity}_i$ . According to the third hypothesis of this paper, the risk level of private entrepreneurs on P2P lending platform is lower than that of the wage earner, reflecting that the interest rate which private entrepreneurs are willing to pay is likely to be lower, namely, the coefficient  $\beta$  of  $\text{identity}_i$  is significantly less than 0. I add the same control variables as used in model (1) to control the impact of borrower's personal characteristics, loan characteristics and

year factors on the borrowing cost.

Focusing on the last question concerned in this paper, that is, whether the investor's discrimination against the borrower's occupational identity is efficient statistical discrimination or inefficient tasted-based discrimination. This paper uses model (4) to test the default situation of borrowers with different occupational identities.

$$\Pr(\text{default}_i = 1) = \phi(\beta * \text{identity}_i + \text{Controls} + \varepsilon_i) \quad (4)$$

Specifically, where  $\text{identity}_i$  is a dummy variable indicating the repayment status of each loan. If the loan is recognized as bad debt by the platform, the value is equal to 1, otherwise it is equal to 0. If coefficient  $\beta$  of  $\text{identity}_i$  is significantly less than 0, it shows that the default rate of borrowers with occupational identity of the private entrepreneur is indeed lower. That is, if investors prefer private entrepreneurs and discriminate against the wage earner, the discrimination or preference is rational, and the discrimination based on the occupational identity is the effective statistical discrimination. If coefficient  $\beta$  of  $\text{identity}_i$  is significantly greater than 0, it means that investors have more discrimination against the wage earner, then this discrimination is irrational, because the discriminated wage earner have a lower default rate. That is to say, this kind of discrimination is likely to be caused only relying on likes and dislikes of investors, and it is inefficient tasted-based discrimination.

## 5. Empirical Results

### 5.1 The impact of occupational identity on the probability of successful funding

Table 3 is the regression result of occupational identity on the probability of successful funding. No control variable is added in the first column; the coefficient of  $\text{identity}_i$  is above zero and is significant at the 1% level. In the second and third columns, the variables on borrowers' attributes and loan characteristics are added while the coefficient of  $\text{identity}_i$  is still significantly above zero. These results prove that occupational identity has an effect on the probability of successful funding, or in other words, occupational identity enables private entrepreneur borrowers to more easily obtain loans successfully. This also shows that investors prefer private entrepreneurs while discriminating against wage earners.

In addition, the results from the second to the third column also show that not only occupational identity influences the probability of successful funding, the credit rating and information verifications of borrowers also significantly influence the probability of successful funding. The higher the credit rating of the borrower is, the higher the probability of successful funding is, which is consistent with people's common sense. The information verifications on P2P platform, particularly such verifications as home address, identity card, work and marriage certificate all exert a positive impact on the probability of successful funding.

However, the effect of some control variables on the probability of successful

funding is inconsistent with common sense or people's intuition. For example, the higher the interest rate is, the lower the probability of successful funding is. This may be attributed to the fact that P2P platform's interest rate for borrowing is far ahead of that of fixed-term deposit of banks and some financial products of some banks. But since P2P platform is an important channel for family financial management, investors invest on P2P platforms mainly for the purpose of obtaining higher return on investment with lower risks. Therefore, a higher interest rate provided by P2P borrowing application means that the default risk for the borrowers is also higher. Some investors may automatically choose to avoid these loan listings, leading to the phenomenon that the interest rate has a negative effect on the probability of successful funding.

**Table 3: Impact of Occupational Identity on Probability of successful funding**

	Success		
	(1)	(2)	(3)
identity	0.330***	0.317***	0.288***
	(42.87)	(30.88)	(26.80)
risk level		0.603***	0.531***
		(130.27)	(108.39)
amount		-0.266***	-0.274***
		(-62.64)	(-62.00)
interest		-0.072***	-0.075***
		(-48.51)	(-48.04)
term		0.006***	0.007***
		(11.15)	(12.00)
male		0.029**	0.038***
		(2.35)	(2.91)
age		0.035***	0.027***
		(58.32)	(40.23)
mobile_cert			-0.452***
			(-29.82)
address_cert			0.448***
			(30.83)
work_cert			0.416***
			(30.68)
network_cert			-0.038
			(-0.93)
idcard_cert			1.293***
			(47.76)
marriage_cert			0.270***
			(27.89)
credit_cert			0.010
			(0.65)
constant	-1.732***	-0.011	-0.928***
	(-299.87)	(-0.26)	(-17.07)
Year fixed effect	Y	Y	Y
Observations	314,646	314,646	314,646
R-squared	0.0686	0.319	0.358

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

## **5.2 Further analysis on discrimination**

In the previous section, it is found in the paper that P2P investors discriminate against the occupational identity of borrowers and such discrimination is directed toward wage earners instead of private entrepreneurs. This may be inconsistent with traditional banking credit, but still makes sense. As explained by the paper when the hypothesis is proposed, these two types of borrowers are likely to differ in credit risks, which may be one of the reasons for the discrimination. In this part, this paper will continue to analyze such discrimination. Firstly, in this paper, borrowers are divided into two groups based on their credit ratings, namely HR-level and non-HR level borrowers. Later, these two groups are tested to see whether investors discriminate against wage earners. It can be seen from the empirical results in Table 4 that for both HR-level and non-HR-level borrowers, investors prefer borrowers with the occupational identity of private entrepreneurs. This indicates that investors' discrimination against wage earners has not changed owing to the difference in credit rating.

**Table 4: Subsample regression results**

	Success			
	HR	HR	Non-HR	Non-HR
identity	0.160***	0.270***	0.139***	0.244***
	(16.30)	(21.88)	(6.56)	(9.55)
amount		-0.310***		-0.182***
		(-59.81)		(-17.54)
interest		-0.077***		-0.095***
		(-43.97)		(-24.54)
term		0.010***		-0.008***
		(16.41)		(-5.57)
male		0.079***		-0.123***
		(5.33)		(-3.71)
age		0.027***		0.014***
		(36.38)		(8.17)
mobile_cert		-0.502***		-0.177***
		(-26.06)		(-6.22)
address_cert		0.689***		-0.021
		(40.80)		(-0.86)
work_cert		0.459***		0.162***
		(29.94)		(4.53)
network_cert		1.137***		0.121***
		(18.91)		(2.85)
idcard_cert		1.531***		-0.204***
		(39.72)		(-2.74)
marriage_cert		0.300***		0.187***
		(27.67)		(7.04)
credit_cert		0.016		0.186***
		(0.93)		(4.52)
constant	-1.849***	-0.422***	-0.113***	2.609***
	(-282.21)	(-6.39)	(-6.01)	(18.05)
Year fixed effect	Y	Y	Y	Y
Observations	298,067	298,067	16,579	16,579
R-squared	0.0211	0.188	0.0622	0.141

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

Next, the paper further examines the moderating effect of discrimination on the relationship between credit rating and probability of successful funding. Column (1) in Table 5 is the regression result without adding the occupational identity variable. Results show that the credit rating exerts a significant effect on the probability of successful funding. Column (2) is the regression result with the addition of the interaction term between credit rating and occupational identity and the coefficient of  $risk\ level_i * identity_i$  is significantly positive. This shows that occupational identity can indeed influence the relationship between credit rating and the probability of successful funding. For borrowers with the occupational identity of private entrepreneurs, one unit of increase in credit rating will lead to faster increase in the probability of successful funding.

**Table 5: Moderating Effect of Discrimination**

	Success		
	(1)	(2)	(3)
identity			0.231*** (13.73)
risk level	0.541*** (110.67)	0.467*** (81.19)	0.511*** (76.65)
risk level*identity		0.136*** (23.20)	0.040*** (4.35)
amount	-0.233*** (-56.72)	-0.263*** (-60.47)	-0.274*** (-62.03)
interest	-0.075*** (-48.38)	-0.075*** (-48.17)	-0.075*** (-48.05)
term	0.003*** (6.24)	0.006*** (10.40)	0.007*** (12.08)
male	0.041*** (3.18)	0.039*** (3.00)	0.038*** (2.91)
age	0.029*** (43.06)	0.027*** (40.72)	0.027*** (40.11)
mobile_cert	-0.461*** (-30.54)	-0.466*** (-30.67)	-0.456*** (-29.97)
address_cert	0.455*** (31.46)	0.455*** (31.37)	0.450*** (30.92)
work_cert	0.401*** (29.74)	0.411*** (30.42)	0.416*** (30.70)
network_cert	-0.043 (-1.04)	-0.003 (-0.08)	-0.028 (-0.68)
idcard_cert	1.302*** (48.62)	1.310*** (48.09)	1.297*** (47.75)
marriage_cert	0.290*** (30.13)	0.281*** (29.14)	0.271*** (28.02)
credit_cert	0.020 (1.32)	0.017 (1.13)	0.011 (0.73)
constant	-1.290*** (-24.60)	-0.945*** (-17.22)	-0.898*** (-16.36)
Year fixed effect	Y	Y	Y
Observations	314,646	314,646	314,646
R-squared	0.354	0.357	0.358

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

This paper also discusses the relationship between occupational status and the interest rate. Through this empirical test, the paper can determine the difference in the overall credit risk status between private entrepreneurs and wage earners from the perspective of interest rate, thus analyzing the possible reasons for investors' discrimination against the occupational identity of investors. The regression result is shown in Table 6. The coefficient of *identity<sub>i</sub>* is significantly negative. This shows that occupational identity further decreases the overall interest rate of borrowers with the occupational identity of private owners. This also excludes a possibility for explaining discrimination. Or rather, the difference in the interest rate of these two types of borrowers may not be the reason for the discrimination.

**Table 6: Relationship between Occupation Identity and Interest Rate**

	Interest Rate		
	(1)	(2)	(3)
identity	-0.285*** (-20.58)	-0.095*** (-6.32)	-0.075*** (-4.98)
risk level		-0.682*** (-75.60)	-0.700*** (-69.75)
amount		-0.006 (-1.12)	-0.002 (-0.28)
term		0.043*** (59.99)	0.043*** (59.99)
male		0.325*** (18.77)	0.321*** (18.58)
age		-0.007*** (-7.10)	-0.003*** (-3.13)
mobile_cert			0.115*** (4.37)
address_cert			0.632*** (24.73)
work_cert			-0.057*** (-3.60)
network_cert			-0.343*** (-4.23)
idcard_cert			-0.192*** (-6.16)
marriage_cert			-0.198*** (-15.07)
credit_cert			-0.279*** (-15.13)
constant	14.002*** (1,545.33)	13.940*** (247.51)	14.202*** (217.65)
Year fixed effect	Y	Y	Y
Observations	314,646	314,646	314,646
R-squared	0.103	0.135	0.138

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

### 5.3 Relationship between occupational identity and the loan default rate

Through previous argument, this paper has proved that investors prefer borrowers with occupational identity as the private entrepreneur, and more discriminate against the wage earner. Next, in view of hypothesis 5, this paper further tests whether such discrimination is efficient statistical discrimination or inefficient tasted-based discrimination. Table 7 shows the empirical results of model (4). After controlling the variables that may affect the default rate, the coefficient of  $identity_i$  is always greater than 0, and is significant at the 1% level. This shows that borrowers with occupational identity as the private entrepreneur have a higher default rate than those as the wage earner. It also shows that although investors prefer private entrepreneurs, this preference is not due to the lower default rate of private entrepreneurs; on the contrary, the default rate of private entrepreneurs is higher and that of the wage earner discriminated against by investors is lower. Therefore, the discrimination of investors is irrational and inefficient tasted-based discrimination.

Then look at the significance of coefficients of the control variables. Borrowers with higher credit rating have the lower default rate. The characteristics including male and age of borrowers can significantly increase the default risk of borrowing. From the perspective of loan characteristics, the longer loan term, the larger amount and the higher interest rate mean the higher possibility of default, which is consistent with the situation in the traditional credit process.

**Table 7: Relationship between occupational identity and the loan default rate**

	Default		
	(1)	(2)	(3)
identity	0.223***	0.338***	0.358***
	(7.98)	(8.36)	(8.50)
risk level		-0.868***	-0.879***
		(-30.80)	(-30.54)
amount		0.147***	0.132***
		(7.46)	(6.38)
interest		0.075***	0.077***
		(11.36)	(11.13)
term		0.028***	0.024***
		(16.22)	(13.48)
male		0.107**	0.099**
		(2.51)	(2.24)
age		0.011***	0.011***
		(5.39)	(4.97)
mobile_cert			0.105**
			(2.37)
address_cert			0.129***
			(3.18)
work_cert			1.394***
			(8.72)
network_cert			0.031
			(0.20)
idcard_cert			0.119
			(0.93)
marriage_cert			-0.015
			(-0.45)
credit_cert			1.028***
			(6.06)
constant	-0.887***	-3.062***	-5.422***
	(-51.32)	(-14.83)	(-16.11)
Year fixed effect	Y	Y	Y
Observations	20,412	20,412	20,412
R-squared	0.0820	0.290	0.328

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

## 6. Robustness Tests

Through previous analysis, the five questions concerned in this paper have already been proved. However, the conclusions may still be interfered by the sample selection factors. In this part, the paper examines the robustness of the results in a number of ways.

### 6.1 Select samples based on the loan amount

In order to avoid the impact of the loan amount on the conclusion of this paper, samples are re-selected according to the amount of each loan. This paper excludes samples with amount exceeding RMB 250,000 and less than RMB 5000 to avoid the influence of samples with too big or too small amount. The first column in Table 8 is the impact of the borrower's occupational identity on the probability of successful funding. It can be seen that the coefficient of  $identity_i$  is still significantly positive, and so is the coefficient of  $risk\ level_i * identity_i$ . This shows that discrimination based on occupational identity also exists in the re-selected samples, and it still has a moderating effect on the relationship between the credit rating and the probability of successful funding. The second column is the regression result of occupational identity to lending rate, and the third column is the regression result of occupational identity to default rate. All these results are consistent with the previous results.

**Table 8: Regression results on reselected samples according to loan amount**

	Success	interest	default
identity	0.278***	-0.094***	0.384***
	(14.70)	(-5.95)	(8.87)
risk level	0.503***	-0.681***	-0.889***
	(63.77)	(-59.20)	(-29.56)
risk level*identity	0.074***		
	(6.89)		
amount	-0.424***	-0.059***	0.087***
	(-72.02)	(-8.44)	(3.80)
interest	-0.102***		0.075***
	(-51.89)		(10.21)
term	0.009***	0.040***	0.024***
	(14.31)	(54.26)	(13.28)
male	0.037**	0.253***	0.108**
	(2.57)	(14.07)	(2.40)
age	0.030***	-0.002**	0.012***
	(40.62)	(-2.12)	(5.24)
mobile_cert	-0.455***	0.118***	0.123***
	(-27.35)	(4.25)	(2.69)
address_cert	0.471***	0.648***	0.122***
	(29.18)	(24.39)	(2.91)
work_cert	0.450***	-0.049***	1.681***
	(29.60)	(-2.92)	(7.75)
network_cert	-0.529***	-0.147	0.170
	(-9.53)	(-1.39)	(1.07)
idcard_cert	1.399***	-0.118***	0.101
	(47.10)	(-3.58)	(0.73)
marriage_cert	0.288***	-0.176***	0.020
	(26.93)	(-12.89)	(0.58)
credit_cert	0.056***	-0.226***	1.334***
	(3.35)	(-11.75)	(5.77)
constant	0.727***	14.744***	-5.605***
	(10.56)	(186.82)	(-13.41)
Year fixed effect	Y	Y	Y
Observations	263,392	263,392	17,415
R-squared	0.377	0.150	0.305

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

## **6.2 Select samples based on borrowers' age**

The second robustness test of this paper is to re-select samples according to the age of borrowers, which aims to avoid the age interference that the sample borrowers who is too young or too old for the conclusion of this paper. The borrower's age in this paper is limited to 25 - 50 years old. The regression results are shown in Table 9, in which the first column is that of the influence of the borrower's occupational identity to the probability of successful funding and its moderating effect, the second column is that of the occupational identity to the interest rate of the borrowing target, and the third column is that of the professional identity to the default rate. These results are consistent with the previous results, showing that the research conclusions are not affected by the samples with borrowers being too old or too young.

**Table 9: Regression results on reselected samples according to borrowers' age**

	Success	interest	default
identity	0.231*** (13.01)	-0.057*** (-3.46)	0.376*** (8.60)
risk level	0.491*** (71.05)	-0.702*** (-67.74)	-0.853*** (-29.32)
risk level*identity	0.045*** (4.72)		
amount	-0.279*** (-59.68)	0.008 (1.39)	0.131*** (6.09)
interest	-0.076*** (-45.38)		0.078*** (10.92)
term	0.007*** (10.84)	0.042*** (53.88)	0.024*** (12.82)
male	0.022 (1.59)	0.281*** (14.43)	0.120*** (2.58)
age	0.023*** (28.27)	-0.002* (-1.66)	0.012*** (4.51)
mobile_cert	-0.451*** (-28.05)	0.108*** (3.76)	0.090* (1.95)
address_cert	0.444*** (29.08)	0.629*** (23.35)	0.119*** (2.82)
work_cert	0.422*** (29.17)	-0.039** (-2.24)	1.361*** (8.39)
network_cert	-0.119*** (-2.70)	-0.347*** (-4.06)	0.147 (0.95)
idcard_cert	1.273*** (44.17)	-0.232*** (-6.58)	0.077 (0.58)
marriage_cert	0.242*** (23.48)	-0.203*** (-14.44)	-0.002 (-0.06)
credit_cert	0.019 (1.20)	-0.302*** (-14.85)	0.992*** (5.73)
constant	-0.627*** (-10.36)	14.140*** (185.72)	-5.393*** (-15.41)
Year fixed effect	Y	Y	Y
Observations	243,392	243,392	18,689
R-squared	0.355	0.142	0.329

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

### **6.3 Random selection of samples**

The third robustness test in this paper is the random selection of 50% samples. In addition to the amount and the age of borrowers discussed in the first two robustness tests, the conclusions of this paper may be influenced by many other factors, such as different years and provinces. That is to say, the conclusions of this study may not be applicable to all P2P borrowers, but may only be applicable to some samples of certain years or provinces. Therefore, in order to avoid the interference of these factors, I randomly selected 50% of the samples to test the previous question. Table 10 is the regression results. From the first to the third columns, we can find that some of the findings in this paper are still valid in randomly selected samples, which shows that the conclusions in this paper are robust.

**Table 10: Regression results using randomly selected samples**

	Success	interest	default
identity	0.247***	-0.074***	0.302***
	(10.45)	(-3.50)	(4.99)
risk level	0.509***	-0.696***	-0.871***
	(54.57)	(-49.31)	(-21.40)
risk level*identity	0.032**		
	(2.54)		
amount	-0.277***	-0.013*	0.139***
	(-44.31)	(-1.69)	(4.73)
interest	-0.074***		0.074***
	(-33.96)		(7.57)
term	0.007***	0.044***	0.023***
	(9.09)	(43.57)	(9.03)
male	0.025	0.314***	0.066
	(1.35)	(12.84)	(1.07)
age	0.027***	-0.002	0.013***
	(28.97)	(-1.35)	(4.13)
mobile_cert	-0.429***	0.114***	0.025
	(-20.04)	(3.09)	(0.40)
address_cert	0.427***	0.651***	0.120**
	(20.61)	(17.95)	(2.07)
work_cert	0.405***	-0.033	1.485***
	(21.28)	(-1.45)	(6.24)
network_cert	-0.062	-0.302***	0.146
	(-1.08)	(-2.67)	(0.67)
idcard_cert	1.302***	-0.181***	0.287
	(33.66)	(-4.13)	(1.52)
marriage_cert	0.274***	-0.209***	-0.027
	(19.99)	(-11.25)	(-0.59)
credit_cert	0.017	-0.289***	1.049***
	(0.80)	(-11.07)	(4.07)
constant	-0.905***	14.246***	-5.676***
	(-11.64)	(154.30)	(-11.33)
Year fixed effect	Y	Y	Y
Observations	157,323	157,323	10,223
R-squared	0.357	0.139	0.323

Note. \*, \*\*, and \*\*\* represent significance at the 10%, 5%, and 1% level, respectively.

## **7. Conclusion**

In industrialized and post-industrialized societies, occupational identity greatly determines people's income levels, and also reflects one's social status to a certain extent. Even many sociologists regard occupational identity as the basis for dividing social classes. At present, discrimination against people with specific occupational status is widespread in various areas of social life, which has become the focus of social issue in the whole society.

Previous studies in the field of credit discrimination mostly focused on gender, racial or regional discrimination, and few explored whether there was discrimination based on the occupational identity of borrowers in the lending process. Using loan data obtained from a leading P2P platform in China, this paper examines whether P2P investors will discriminate against borrowers with certain occupational identities. The results show a different phenomenon from common sense or traditional bank lending, that is, P2P investors prefer private entrepreneurs and discriminate against the salaried class borrowers. This paper believes that this phenomenon may be caused by the high risk level of borrowers of P2P lending platforms and the possible adverse selection problem, which leads to the higher overall credit rating of private entrepreneurs than that of the salary earners. However, after examining the default rate of these two groups, it is found that the salary earners have a lower default rate than private entrepreneurs, which shows that the discrimination of P2P investors against the salary earners is irrational, that is, this kind of discrimination is the inefficient tasted-based discrimination.

This paper extends the research on the behavior of P2P lending investors by using the loan data of micro-individuals from an online lending platform, providing new empirical evidence for understanding the behavior of P2P investors. However, although the results directly prove that there is discrimination against certain occupational identity of borrowers in the process of P2P lending, because the data used in this paper are from the early stage of the P2P development, the investment behavior of P2P investors is likely to change over time, such as learning from previous investment experiences and lessons. Therefore, whether this kind of discrimination will exist for a long time is a further study in the follow-up of this paper.

## **References**

- [1] A. F. Alesina, F. Lotti and P. E. Mistrulli, Do women pay more for credit? evidence from Italy, *Journal of the European Economic Association*, 11(s1), (2013), 45-66.
- [2] K. Arrow, Higher Education as a Filter, *Journal of Public Economics*, 2(3), (1973), 193-216.

- [3] A. Bellucci, A. Borisov and A. Zazzaro, Does gender matter in bank–firm relationships? Evidence from small business lending, *Journal of Banking & Finance*, 34(12), (2010), 2968-2984.
- [4] M. Billig and H. Tajfel, Social categorization and similarity in intergroup behavior, *European Journal of Social Psychology*, 3(1), (1973), 27-52.
- [5] L. Blanchard, B. Zhao and J. Yinger, Do lenders discriminate against minority and woman entrepreneurs? *Journal of Urban Economics*, 63(2), (2008), 0-497.
- [6] D. G. Blanchflower, P. B. Levine and D. J. Zimmerman, Discrimination in the small-business credit market, *Review of Economics and Statistics*, 85(4), (2003), 930-943.
- [7] L. Brandt and H. Li, Bank discrimination in transition economies: ideology, information, or incentives? *Journal of Comparative Economics*, 31(3), (2003), 387-413.
- [8] K. S. Cavalluzzo and L. C. Cavalluzzo, Market structure and discrimination: The case of small businesses, *Journal of Money, Credit, and Banking*, 30(4), (1998), 771-792.
- [9] X. Chen, L. Zhou and D. Wan, Group Social Capital and Lending Outcomes in the Financial Credit Market: An Empirical Study of Online Peer-to-Peer Lending, *Electronic Commerce Research and Applications*, 15, (2016), 1-13.
- [10] S. Claessens, E. Feijen and L. Laeven, Political connections and preferential access to finance: the role of campaign contributions, *Journal of Financial Economics*, 88(3), (2008), 554-580.
- [11] A. Cozarenco and A. Szafarz, Gender biases in bank lending: Lessons from microcredit in France, *Journal of Business Ethics*, 147(3), (2018), 631-650.
- [12] G. Dorfleitner, C. Priberny, S. Schuster, J. Stoiber, M. Weber, I. De Castro and J. Kammler, Description-text related soft information in peer-to-peer lending - evidence from two leading european platforms, *Journal of Banking & Finance*, 64(3), (2016), 169-187.
- [13] M. Herzenstein, S. Sonenshein and U. M. Dholakia, Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-Peer Lending Decisions, *Journal of Marketing Research*, 48(SPL), (2011), S138-S149.
- [14] H. Li, L. Meng, Q. Wang and L. A. Zhou, Political connections, financing and firm performance: evidence from chinese private firms, *Journal of Development Economics*, 87(2), (2008), 283-299.
- [15] M. Lin and S. Viswanathan, Home bias in online investments: an empirical study of an online crowd funding market, *Management Science*, 62(5), (2016), 1393-1414.
- [16] M. Lin, N. R. Prabhala and S. Viswanathan, Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending, *Management Science*, 59(1), (2013), 17-35.
- [17] E. Phelps, The Statistical Theory of Racism and Sexism, *American Economic Review*, 62(4), (1972), 659-661.