Retraction

Retracted: Evaluating Investors' Recognition Abilities for Risk and Profit in Online Loan Markets Using Nonlinear Models and Financial Big Data

Journal of Function Spaces

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. He, P. Xia, B. Li, and J. Liu, "Evaluating Investors' Recognition Abilities for Risk and Profit in Online Loan Markets Using Nonlinear Models and Financial Big Data," *Journal of Function Spaces*, vol. 2021, Article ID 5178970, 15 pages, 2021.

Research Article

Evaluating Investors' Recognition Abilities for Risk and Profit in Online Loan Markets Using Nonlinear Models and Financial Big Data

Qizhi He [ID],[1] Pingfan Xia [ID],[2] Bo Li,[2] and Jia-Bao Liu [ID][3]

[1]*School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China*
[2]*School of Finance, Anhui University of Finance & Economics, Bengbu, 233030 Anhui, China*
[3]*School of Mathematics and Physics, Anhui Jianzhu University, Hefei, 230601 Anhui, China*

Correspondence should be addressed to Qizhi He; happyhefei2000@163.com

Financial big data are obtained by web crawler, and investors' recognition abilities for risk and profit in online loan markets are researched using heteroskedastic Probit models. The conclusions are obtained as follows: First, the preference for the item is reflected directly in the time and indirectly in the number of participants for being full, and the larger the preference, the shorter the time and the fewer the participants. Second, investors can discriminate the default risk not reflected by the interest rate, and the bigger the default risk, the longer the time and the more participants being full. Third, investors can discriminate the pure return rate deducted from the maturity term and credit risk, and the higher the return, the shorter the time and the fewer the participants being full. Fourth, default risks are reflected well by online loan platform interest rates, and inventors do not choose the item blindly according to the interest rate but consider comprehensively the profit and the risk. In the future, interest rate liberalization should be deepened, the choosing function of interest rates should be played better, and the information disclosure, investor education, and investor effective usage of other information should be strengthened.

1. Background

With the rapid development of electronic information, internet finance, which has low transaction costs, low participation threshold, and convenient features, gets development in full swing. But there are also many internet finance platforms that have gone bankrupt in recent years. In this context, the government pays more attention to internet finance, and the focus also changes from healthy and standardized development to preventing accumulated risks and strengthening supervision. Thus, it is necessary to research issues of yield and risk of internet finance and to ascertain how investors can participate in internet finance platforms and whether they can discriminate risk and yield.

Traditional debit and credit usually take bank as the intermedium, and depositors and lenders are passive receivers of the interest rate, and thus, characteristics of rate marketization cannot be reflected completely. Although the online-loan-platform interest rate is also one kind of non-fully market-oriented interest rate, the full bid rate, the result of depositors and lenders weighing each other, reflects well the characteristics of interest rate marketization. Traditional financial institutions such as banks do not announce the information of depositors and lenders, but online-loan-platforms announce information of borrowers to potential lenders to promote a deal. Online loan platforms have more characteristics of Financial Big Data than traditional financial institutions and thus can provide rich data resources for researching how online investors weight risks and rewards of online debt items and whether they can.

The online loan platform, without the participation of traditional banks and other financial institutions, can reduce costs and improve the efficiency of capital allocation through direct financing between borrowers and lenders and thus affords a new path to solve the problems of difficult and expensive financing for Small and Medium-Sized

Enterprises (SMEs). The online loan platform, as a new financial medium, broadens investing and financing channels of grassroot debit and credit and makes the common people have a chance to get a higher return. As to whether the online loan platforms have a price discovery function, whether investors are rational and can recognize risks and returns, and whether investors can effectively identify different default risks behind the same interest rate and the net yield difference after deducting maturity period and credit risk, all those need research deeply, and there are great theoretical and practical meanings for understanding and regulating behaviors of Chinese online loan investors and promoting online loan developing healthily. The online loan is an important aspect of internet finances, and the failure probability is very high in recent years, but there are many fake internet finances which are excluded by the paper. The paper researches the recognition ability of risk and return and expects to afford suggestions for formal online loan's healthy and sustainable development based on the Renrendai Online Loan Platform.

2. Literature Reviews

Big financial data accumulated by online loan provide material for deeply researching references and behavior of borrowers and investors. Existing research mainly focuses on researching the full rate, the default rate, and investor's rational consciousness and behavior, and these issues mix up and can be roughly divided into the following categories by evolving process.

Factors influencing the full rate and lender's judgment have been researched by many literatures, and the full rate is a fundamental problem about online loan. Klafft [1] has tested the factors influencing the success of online loans based on the America Prosper platform, and the results show that credit rating, individual character, etc. are important to the success rate of the item. Li et al. [2] analyze the basic statistical characteristics of the online loan item using the data of ppdai and find that basic information of the borrower and item has an important effect on the loan success rate. Li et al. [3] find that descriptive information having a positive impact on the full rate and the more positive information are beneficial to successful fundraising. Liu et al. [4] research lenders' decision-making characteristics and found that friendship has an important effect in the online loan market, and there is a herding effect in the market, namely, lenders following their friends' lending decision. Wan et al. [5] find that the initial trust and consciousness on yield are the main factors affecting lenders' lending in the online loan market.

Some researchers are concerned about borrowers' final activities, namely, factors influencing the default rate, and some are focused on whether principal and interest of online loan can be paid on schedule. Iyer et al. [6] test empirically the role of credit score to online loan's default rate based on data of Prosper, and the results showed that credit level has significant influence on the default rate. Liao et al. [7] test empirically the relationship between interest rates and default rates using data from the Renrendai website.

Serrano-Cinca et al. [8] research empirically factors influencing online loan default rates based on data from the Lending Club which is the biggest P2P company in the USA. Emekter et al. [9] research the characteristic of the P2P online loan using data from the Lending Club website, and the empirical test showed that indicators such as credit rating have a significant effect on default rates of items and that high interest rate corresponding to high risk cannot compensate the higher loan default rate. Ge et al. [10] test empirically the influence of social medium information shown by the borrower itself to the online loan default rate. Liu et al. [11] test empirically the forecasting effect of lender's information on default rates of items using data from the Renrendai website.

There are also many literatures that combined the former two questions and researched comprehensively success rates and default rates of fundraising. Using data from the Prosper company, Ravina [12] finds that borrower's individual characteristics such as ethnicity, looking credible, beauty, and body weight have an important effect on financing success rate, but those characteristics except beauty have no influence on the late performance of the item, and although beautiful borrowers more easily get a loan from the online market, they are more likely to delay repayments. Freedman and Jin [13] test empirically the information discrimination of online loans and thought that there are three issues: adverse selection, lender misjudgment, and high interest rate corresponding to high risk in the online loan market, and the former two issues are unique to the online loan market and can be relieved by announcing more borrowers' information and lenders' studies, and the last issue is also existing in traditional markets, and online loans will eventually compete with traditional banks directly. Guo [14] researched roles of internet nicknames and real names on fundraising success rate and default rate, and the empirical test shows that real names cannot increase the success rate and decrease the default rate. Yue et al. [15] test empirically market information's role on investor behavior and forecasting item's default rate. Guo [16] researched the role of marriage in online loans and found that marriage benefits for both increasing fundraising success rate and decreasing default rate. Zhang and Cai [17] research "title bias," namely, the difference in the role of the title to the full rate and default rate. Xu and Chau [18] research the role of communication between lender and borrower on the full rate and default rate and showed that information communication plays a significance effect on the full rate but a nonsignificant effect on the default rate. Caldieraro et al. [19] research the role of nonverified information offered by the borrower to fundraising and the item's late performance and found that verified and nonverified information both have an important effect on online loans. Hu et al. [20] research the performance, financing difficulty, and financing cost of the peasantry and low-income people on the Renrendai platform by combining inclusive finance and online loans. Babaei and Bamdad [21] have evaluated the return and risk of the P2P item using Artificial Neural Network and Logistic function, respectively.

With the deepening of research, some scholars have done deep and detailed research on investors' rational consciousness and choosing behavior. Freedman and Jin [22] show that studying benefits clearing information asymmetry among market participants, and online loan lenders step away from high-risk items and subprime borrowers are excluded from the Prosper website.

Liao et al. [7] research choosing behaviors of online loan investors on the background of China's non-completely-market-oriented interest rate and tested whether information in addition to interest rate has an indicative effect, and investors can discriminate the different default risk implied in the same interest rate based on testing the relationship between interest and default rates, but they have not considered the time value of the interest rate and excluded the credit risk from the interest rate. Gao et al. [23] test "gender effect" using data from the Renrendai website and showed that men operate more frequently based on self-confidence and have a lower yield because of exchanging costs than women in the online loan market. Dorfleitner et al. [24] have researched the role of credit risk and social impact on interest-free P2P lending using Logistic and Tobit models.

Hu and Song [25] test well the investor's rational consciousness from two angles which are the default rate and the full rate and showed that Chinese online investors have rational consciousness of preferring yield and avoiding risk, but they have not tested the relationship between investor number and default rate at full circumstance, and they discriminated yield and default risk by choosing different kinds of items. Although it is rational relatively, there are further chances in the interaction of interest rate and default rate. In the meantime, Liao et al. [7] and Hu and Song [25] have not considered the heteroskedasticity effect of the model. The paper tests the investors' discriminating ability of yield and risk after considering time value of interest rate, relationship between interest rate and credit risk, and heteroskedasticity effect.

Existing research focuses mainly on the full rate and the default rate, and some discuss deep-seated questions such as investors' rational consciousness, but they are not deep enough, and methods used are relatively simple. In the meantime, information asymmetry and adverse selection are more serious in online loan markets than traditional offline markets: on the one hand, online loan lenders have difficulties to get complete credit notes of borrowers; on the other hand, many online loan borrowers are fundraisers who have difficulties in getting credit debts offline [22]. Unlike American market-oriented completely interest rate, the Chinese online loan interest rate is incompletely market-oriented and set up initially by borrowers according to self-conditions within ranges specified by the government and then bidden by lenders according to items' interest rates and information. It is to be called a full bid if the investment fund reaches the amount the borrower is planning to get, and the corresponding interest rate is effective; otherwise, it is to be called a flow bid, and the corresponding interest rate is noneffective. Compared with traditional bank credit debts, lenders in the online loan market are at information disadvantages, and they can only decide based on the items

and borrowers' information published by the website and their own experience and then whether the decision is rational, but can the investor discriminate the default rate difference behind the same interest rate? Can the investor discriminate the yield difference behind the same credit risk? There are great theoretical and real meanings for sorting out these questions in the background of constant advancement and coming to an end of the market-oriented Chinese interest rate.

The work of the paper is mainly exhibited: First, the rational consciousness of the online loan investor has been researched from two angles which are the interest rate and the default risk. Interest rate and default risk are influencing each other, but most of the exiting research have not eliminated the mutual influence when testing the impact of interest rate and default rate on investor behavior. And then on the one hand, there may be implied the impact of default risk when testing the impact of interest rate on investor behavior; on the other hand, there may be implied the impact of interest compensation when testing the impact of default risk on investor behavior. (Liao et al. [7] research the impact of default risks not reflected by interest rates on investors' behaviors, and they deduct the implied effect of interest rate on default risk, but they do not deduct the implied effect of default rate on interest rate, namely, the influence of interest rate on investor's behavior maybe is caused by the default risk corresponding to the interest rate. And they have not considered the term structure of the interest rate and heteroskedasticity effect.) The paper not only excludes the role of interest rate from default risk but also excludes the role of default risk from interest rate and tests impacts of default risk not reflected by interest rate and pure yield excluded time term structure and default risk on investor behavior. Second, the paper is based on micro-data, and macro and micro are combined. The paper crawled more than 0.3 million data from the Renrendai website using web crawling technology, and the data reflect in detail the microcharacteristics of online loans, and thus, effectiveness, credibility, and reality microfoundation of empirical tests are ensured. Term structure theory of interest rates in the macrofield is applied to the microfield, and the influence of different maturities on the interest rate is considered. Third, at the research perspective, both borrowers' characteristics and lenders' subjective initiatives are considered, and lenders' identification ability of yield and default risk is researched, and at the econometric model, heteroskedasticity is considered, characteristics of Chinese online loans are analyzed, and suggestions on how to develop healthy online loans are offered based on situation analysis, theory combing, and empirical tests.

3. Theoretical Analysis and Research Hypothesis

3.1. Theoretical Analysis. Whether traditional offline or online financial market, investors mainly think of two factors: yield and risk, and for online loan platforms, investors' concern is mainly on the item's interest rate and default risk. Because we research the default risk (credit risk) directly,

investors avoiding risk is in keeping with facts. Investors chasing yield and avoiding default risk has become a consistent conclusion. Liao et al. [7] showed that investors in the Renrendai Market are disgusted with risk and thus chase the minimized risk at the equal yield. Hu and Song [25] test the phenomenon that investors of online loans prefer the item with lower default risk at the same yield or higher yield with the same default risk. Next, we should research further how to measure the default risk difference behind the same interest rate, the yield difference behind the same default risk, and whether investors can determine the difference and how to determine it. Akerlof [26] discusses commodity quality uncertainty, information asymmetry, and market structure. In some markets, consumers evaluate utilities of potential purchases according to market statics information, and this will produce the profit difference between the whole and the individual seller. Benefits for all parties can be enhanced by government regulation. Unwritten promises are preconditions for many products and trades that proceed well, but adverse choices caused by information asymmetry may make prices disorderly and the configuration efficiency low. Borrowers in online markets are at an information advantage, and investors can only judge whether borrowers can pay back capital and interest timely in the future according to items and borrowers' information published by borrowers in online platforms and investors' experience. In an imperfect market, different dealers hold different information, some hold specific information and some do not hold specific information only aimed at communication, and some may judge erroneously public information, and generally, dealers who hold advantageous information will gain [27]. People's attention is limited, and mutual interference will appear when attention increase cannot satisfy increased needs [28], and online market platforms publish much information on items and borrowers, and the information mingles with each other, and investors should pay consistent attention to this information and give correct judgments. Peng and Xiong [29] analyze investors' classification learning abilities and their roles in asset dynamic pricing; investors prefer applying limited attention to classification learning and are adept in using markets and industry information but are weak on company-specific information, and thus, investors in online platforms also prefer classification recognition and learning according to information published by the platform and then give investment decision.

In a prefect nonarbitrage market with transparent information, the yield and the risk correspond to each other. But as said in the former, online market information is asymmetric, and risk cannot be indicated by yield and yield cannot be indicated by risk. Default risks behind the two items with the same interest rate have heterogeneity; whether investors can and how to discriminate the heterogeneity need to be discussed further. With the gradual deepening of interest rate marketization, Chinese interest rate marketization has been to the final stage. Online loan interest rate is marketized incompletely, and borrowers decide interest rate levels themselves in the range specified by the government, and lenders decide whether to invest and how much to invest according to information such as borrowers' credit level published in the online website. There are many participants in the online loan website, and information communicates rapidly, and thus, there are generally multiple participants involved in the full item finished in the specified time, and this embodies fully the strength of the market. Investors make decisions mainly considering from the two angles of yield and risk.

In short, theories and existing empirical research show that rational investors in online loan markets chase default risk minimizing at the same yield or yield maximizing at the same risk. The default risk behind the same interest rate and interest rate corresponding to the same default risk may be different, and then, whether Chinese online loan investors can effectively discriminate these differences, and how to behave if they can, all these need to be tested by empirical data.

3.2. Research Hypothesis. The former analyses show that default risk behind the same interest rate in online loan markets may be different, and there may exist a default risk not reflected by the interest rate, which is later called the excess default risk. The interest rate level corresponding to the same default risk may be also different, and there may exist an interest rate that does not correspond to the default risk, which is later called the short excess yield.

Investors who can discriminate the kinds of excess default risk and excess yield will show different preferences to the corresponding item. Generally, the more investors prefer the item, the more investors will make investing decisions rapidly and the investing amount, and thus, items preferred by investors will be full in shorter time and need fewer people, and thus, the preference of investors for the item can be reflected by the full time and full participant number. Rational investors are prone to avoid items with extra default risk, and this will be expressed in two aspects: one is less investors choosing the item and the other is the investor of the item investing with a smaller amount for prudent goal, and the two behaviors will make the full time take longer. Based on the above analysis, we conclude hypothesis 1.

Hypothesis 1 (extra default risk and full time have a positive relationship). At a given yield, the greater the risk uncompensated by the interest rate, the longer the time needed for full.

Not only is the risk not reflected by the interest rate but also the net yield influences investor behavior, and the higher the net yield after being deducted the credit risk and time value, the more investors prefer the item and are willing to invest in the item. Rational investors prefer to choose items with extra yield, and we can get hypothesis 2 similarly to the above analysis.

Hypothesis 2 (extra yield and full time have a reverse relationship). At a given default rate, the greater the net yield after being deducted the credit risk and time value, the shorter the time needed for full.

The preference an investor pays to the item exhibits not only directly in the full time but also indirectly in the participant number for full.

When investors prefer to avoid the item with extra default risk, participant number will appear in two situations: one is that the participant number willing to invest in the item decreases, which makes the item fail, and the other is that at full circumstance, the amount the investor is willing to invest decreases, which leads to the number needed for full to increase. The two situations correspond to two opinions on participant number: one is that the more the public prefers the item, the participant number is more and leads to it easily being full [25], and the other is at full circumstance, the risk not reflected by the interest rate is bigger, the investor is more careful and thus full need the more participant number. And thus, the impact of the investor choosing behavior on the participant number has not coincided, and it is necessary to test empirically further for determining the influencing mechanic and effect. Thus, the impact of investor choosing behavior on participant number has not coincided, and further empirical test is needed to determine the influencing mechanism and result. Because we research the situation of full, it is expected to be the second situation in the paper. Namely, risk premium not reflected by the interest rate has a positive relationship with time and participant number for full. The bigger the risk not predicted by the interest rate exhibits directly in the longer time and indirectly, the more the participant number needed for being full. And thus, we get hypothesis 3.

Hypothesis 3 (extra default risk and participant number for being full have a positive relationship). At a given yield, the bigger the risk not compensated by the interest rate, the more the participant number for being full.

Not only is the risk not reflected by the interest rate but also the net yield has an influence on investor's behavior. The higher the net yield, which is deducted the credit risk and time value, the investor prefers the item more and thus is more willing to invest in it. The higher net yield means fewer participants can complete the full bid, and thus, the smaller the participant number needed for being full. Rational investors prefer to choose items with extra yield, and we can get hypothesis 4 similarly to the above analysis.

Hypothesis 4 (extra yield and participant number for being full have a reverse relationship). At a given default rate, the greater the net yield after being deducted the credit risk and time value, the fewer participants for being full.

4. Research Design, Variable Selection, and Sample Characteristics

4.1. Research Design. The interest rate is one of the most important factors in P2P Lending [24]. Generally, the item's interest rate level corresponds to its risk level, and they interact with each other. If empirical tests directly use the interest rate and default rate of the market, investors' discriminating abilities on risk premium not being reflected by yield and

pure yield after the term structure and credit risk have been deducted cannot be measured effectively, because the interacting effect of yield and credit risk is not discriminated. And thus, we conduct, respectively, the default rate not reflected by yield and the pure yield after the term structure and credit risk are deducted based on the relationships among the default rate, interest rate, and credit level.

$$P(d_i = 1 | r_i) = f(r_i) + e_{1,i}, \tag{1}$$

$$P(d_i = 1 | r_i, x_i) = f(r_t, x_i) + e_{2,i}, \tag{2}$$

$$\Delta P_{d_i} = P(d_i = 1 | r_i, x_i) - P(d_i = 1 | r_i). \tag{3}$$

d_i expresses whether default, r_i expresses interest rate, and P_{d_i} expresses default probity, equation (1) expresses default probity calculated only based on interest rate, equation (2) expresses default probity calculated based on interest rate and other factors which may influence the default rate, and equation (3) measures the default rate which is not reflected by interest rate, measuring the default difference behind the same interest rate, namely, extra-default rate. The smart investor can discriminate the difference, and it is reflected in whether to bid and the bid amount, and in addition directly reflected the time and indirectly participant number needed for being full. To discriminate different maturity influences, term structure theory of interest rate is applied to online loan interest rate, and the interest rates are converted to continuous compound interest. After discriminating impacts of maturity, interest rates are influenced mainly by credit risk, and each item's credit risk corresponds to a level of interest rate, and the gap between it and the factual interest rate measures the yield no credit risk is corresponding to, namely, the extra yield.

$$\Delta r = \frac{\text{Ln}(1 + t * r_t)}{t} - f\left(\frac{\text{Ln}(1 + t * r_t)/t}{\text{cl, hb, ho}}\right). \tag{4}$$

r_t is the interest rate of term t, $(\text{Ln}(1 + t * r_t))/t$ eliminates the influence of the maturity term and translates r_t to the continuous compound interest, $f((\text{Ln}(1 + t * r_t)/t)/\text{cl, hb, ho})$ is the continuous compound interest corresponding to and predicted by the known credit level, and Δr measures the pure yield that being eliminated the differences of maturity term and credit level, namely, the extra yield.

Extra risk and extra yield can be obtained by formulas (3) and (4), and in addition, equations can be made up by taking the time for being full or the participant number for being full as the explained variable. And specific influencing effects of extra risk and extra yield on the time and the participant number for being full can be tested empirically.

4.2. Variable Selection and Sample Characteristics. Different from the completely marketized mechanism of the online loan interest rate decided by the relationship between borrowers and lenders in USA and England, etc., the Chinese online loan interest rate is determined mainly by borrowers according to their situations and is verified by the platform, and then, potential lenders decide whether to invest and the

investment amounts according to information published by the platform. The item fails if the amount investors are willing to invest cannot reach the amount planned to be raised in the given term, which is called out of bid, and the item succeeds if the amount investors are willing to invest reaches the amount planned to be raised in the given term, which is called the full bid. Investors cannot decide the item's interest rate level directly, but they can impact indirectly the final actual interest rate by the model of "vote with feet," and this is an incompletely marketized mechanism of interest rate. Our primary purpose is to research whether investors can discriminate the interest rate and risk of an online loan item, especially the different default risk behind the same interest rate and the different yield behind the same credit risk.

According to the former theoretical analysis, research hypothesis, existing research practices, and data available, we will take the following variables: (1) Whether default (d): it will be assigned 0 if the full item does not default at maturity and 1 if it defaults. (2) Interest rate (r): the promised interest rate of the item. (3) Time for being full (ft): the duration needed for being full, and the unit is day. (4) Number for being full (n): participant number when being full. (5) Total amount (ta): planning to raise the total loan amount, and the unit is yuan. (6) Deadline for repayment (dd): the planned deadline for paying back the raised fund of the item, and the unit is month. (7) Credit level (cl): borrower's credit level, and Renrendai platform offers a comprehensive credit valuing index according to the borrower's various indicators, and seven grades from high to low according to credit level are given—AA, A, B, C, D, E, and HR—and are expressed, respectively, by 7, 6, 5, 4, 3, 2, and 1, and thus, the bigger cl means the higher credit level. (8) Historical borrowing times (hb): the times when the borrower issued a financing project in the Renrendai platform. (9) Historical overdue times (ho): the times when the borrower borrowed successfully and overdue in the Renrendai platform. (10) Year (y): the age of the borrower, and it is required to be from 22 to 70 in the Renrendai platform. (11) Education (e): the borrower's educational background, and it is divided into four grades—high school (or below), college, undergraduate, and postgraduate (above)—is and expressed, respectively, by 0, 1, 2, and 3, and thus, the bigger e means the higher the educational background. (12) Marriage (m): The borrower's marriage situation—1 if married and 0 if not married. (13) Housing (h): the borrower's housing situation—1 if having house and 0 if else. (14) Car (c): the borrower's car situation—1 if having a car and 0 if else. (15) Income (i): the total borrower family income, and the unit is yuan.

In the above, whether default (d) and interest rate (r) are the most concerned indexes by investors, and investors chase high yield and low credit risk, in which generally, the two cannot be both obtained. Time for being full (ft) and number for being full (n) mainly measure investors' preferences to items. Total amount (ta) and deadline for repayment (dd) reflect the basic information of items, credit level (cl), and historical borrowing times (hb), and historical overdue times (ho) reflect borrowers' credit situations. Year (y), education (e), marriage (m), housing (h), car (c), and income (i) reflect borrowers' individual and family situations.

The data are grabbed using web crawl technology and Python3.6 after registering in the Renrendai website which is an online loan platform offering information on loan items. Default or not need not only have been full but also have been finished, and generally, the longest deadline for repayment in the Renrendai platform is 3 years. The sample term we choose is from 2010.10.11 to 2015.01.04, and thus, our entire sample does not conclude the item in repayment. We first grab more than 300,000 sample data and then further remove samples which are not full or are incompatible to the age requirement of the Renrendai website or samples having other obviously abnormal features or missing information, and finally, there are 99,492 samples that can be used.

According to Table 1 at full circumstance, the means of default rates and interest rates are about 5.62% and 12.68%, and the average time and number for being full are about 0.6 hours and 43 people. The average borrowing amount and deadline for repayment are about 56,000 yuan and 2 years. The average credit level is relatively high, historical borrowing times are over 2, and historical overdue times are less than 1. The average age of borrowers is about 39 years, and most borrowers' education gradations are not high. Most borrowers have been married, half of families have a house, and 27.16% of families have a car. Intuitively, the default rate of the Renrendai online loan is relatively low, the interest rate is much higher than the bank interest rate but also in the government-given range, the average time for being full is relatively short and less than 1 hour, and fundraisers who have a historical borrowing record and less overdue are easy to raise funds successfully and be full.

5. Empirical Results

5.1. Measuring Default Rates Not Reflected by Interest Rates.
The most fundamental purpose of lenders investing in items on the online platform is to gain yield, and they are concerned mostly with the default rate and interest rate, and they want to gain a stable high yield and simultaneously fear borrowers defaulting and not repaying the principal and interest on time. In circumstances of nonarbitrage and information transparency, the default rate and interest rate have a linear correspondence relationship, and thus, default levels can be reflected by only using interest rates. We construct the single-factor heteroskedasticity Probit model by taking the default rate as an explained variable and the interest rate as an explanatory variable, and the heteroskedasticity is in the form of [30]

$$P_{d_i} = P(d_i = 1 \mid r) = \Phi(r_i) + \varepsilon_i, \varepsilon_i \sim N(0, \sigma_i^2),$$
$$\sigma_i^2 = e^{2f(r_i)}. \tag{5}$$

Parameters can be estimated by the maximum likelihood function [30]:

$$\log L = \frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n} 2f(r_i) - \frac{1}{2} \sum_{i=1}^{n} e^{-2f(r_i)} \left(P_{d_i} - \Phi(r_i)\right)^2. \tag{6}$$

TABLE 1: Samples of statistical characteristics.

Variables	Sample number	Mean	Standard deviation	Minimum	Maximum
d	99492	0.0562	0.2304	0.0000	1.0000
r	99492	12.6786	1.3382	3.0000	24.4000
ft	99492	0.6076	2.0330	0.0000	36.9083
n	99492	42.6097	59.8358	1.0000	1840.0000
ta	99492	56242.8200	57495.49	3000.0000	3000000.0000
dd	99492	25.1656	10.8038	1.0000	36.0000
cl	99492	5.4630	1.4584	1.0000	7.0000
hb	99492	2.1502	6.9564	1.0000	148.0000
ho	99492	0.5926	2.8495	0.0000	54.0000
y	99492	39.2862	8.2369	22.0000	60.0000
e	99492	0.9819	0.7374	0.0000	3.0000
m	99492	0.7193	0.4493	0.0000	1.0000
h	99492	0.4789	0.4996	0.0000	1.0000
c	99492	0.2716	0.4448	0.0000	1.0000
i	99492	15677.0400	14727.5000	1000.0000	50000.0000

The statistics for heteroskedasticity obey the chi-square distribution [30]:

$$n \log \sum_{i=1}^{n} \frac{\left(P_{d_i} - \Phi(r_i)\right)^2}{n} - \sum_{i=1}^{n} 2\widehat{f}(r_i) \sim \chi^2(1). \quad (7)$$

According to Table 2, first, there is the heteroskedasticity effect. Heteroskedasticity testing shows that the null hypothesis there does not exist and heteroskedasticity is denied, and this illustrates that the traditional Probit model cannot reflect effectively the heteroskedasticity effect that existed in the model, and the heteroskedasticity Probit model needs to be used. Second, interest rates have a significant positive role on default rates. The higher interest rate corresponds to the higher default rate, and this indicates that online loan investors cannot purely choose the item with the high interest rate but must balance default rates and interest rates. Third, the default rate level reflected by the interest rate can be calculated based on formula (3) and the parameters given in Table 2.

If in a complete market-oriented interest rate circumstance as mentioned above, then the interest rate can reflect completely the default risk. In fact, the Chinese online loan interest rate is incomplete market-oriented, it is decided by the borrower unilaterally, and the investor can only take the passive decision mode of "voting with feet." Thus, we add other variables into the default rate estimating model and construct the following multiple-factor heteroskedasticity Probit model, and the heteroskedasticity takes the form of [30]

$$P_{d_i} = P(d_i = 1 \mid r_i, x_i) = \Phi(r_i, x_i) + \varepsilon_i, \varepsilon_i \sim N\left(0, \sigma_i^2\right),$$
$$\sigma_i^2 = e^{2f(r_i, x_i)}. \quad (8)$$

Parameters can be estimated by the maximum likelihood function [30]:

$$\log L = \frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n} 2f(r_i, x_i) - \frac{1}{2} \sum_{i=1}^{n} e^{-2f(r_i, x_i)} \left(P_{d_i} - \Phi(r_i, x_i)\right)^2. \quad (9)$$

The statistics for heteroskedasticity obey the chi-square distribution [30]:

$$n \log \sum_{i=1}^{n} \frac{\left(P_{d_i} - \Phi(_i r, x_i)\right)^2}{n} - \sum_{i=1}^{n} 2\widehat{f}(r_i, x_i) \sim \chi^2(m). \quad (10)$$

According to Table 3, first, there is a heteroskedasticity effect in the driving factor model of the default rate. The heteroskedasticity existing test shows that the null hypothesis that heteroskedasticity does not exist is denied, and this illustrates that the traditional Probit model cannot reflect effectively the heteroskedasticity effect existing in the model, and thus, the heteroskedasticity Probit model needs to be used. Second, total loan amount, deadline, age, and historical overdue times have a significant positive relationship with default rates. The bigger the loan amounts, the bigger the stress for repaying principal and interest in a timely manner, and thus, the bigger the corresponding default rate. The longer the deadline, the larger the uncertainty factor and the bigger the risk for lenders regaining their principal and interest at maturity and thus the corresponding bigger default rate. The older borrowers correspond to the higher default rate possibly because older borrowers in the online loan market have been given the restricted employment chance and reduced ability to repay principal and interest in a timely manner, and thus a corresponding higher default rate. The bigger historical overdue times indicate the

TABLE 2: Single factor estimated results of heteroskedasticity Probit model of default.

Variables		Explained variable: whether default
	dy/dx	z-statistics and p value
r	0.0145	28.0700 ($\leq$0.001)***
Testing whether existing heteroskedasticity		
Null hypothesis	Lr-statistics and p value	Conclusions
Homoskedasticity	104.48 ($\leq$0.001)***	Rejecting null hypothesis, namely, existing heteroskedasticity
Variance equation		
Variables	Coefficient estimated values	z-statistics and p value
r	-0.0367	-11.51 ($\leq$0.001)***

Note: *** means that the corresponding coefficient is significant at 1% significance level.

borrowers having a relatively poor historical performance record, and the significant positive effect of historical overdue times on default rates indicates that the borrowers' repaying behaviors have inertial characteristics and the borrowers' past performance record has a significant forecasting effect on his future repaying behavior. Total loan amount, deadline, age, and historical overdue times that have a significant positive effect on default rates is not only in accordance with theories but also actual situations. Third, credit level, education, historical borrowing times, income, housing, and car have significant negative effects on default rates. Credit level is a comprehensive valuing index given by the online loan platform according to borrowers' various indexes, and the borrower's higher credit level means the borrower's better credibility and the stronger ability and willingness for borrowers to repay principal and interest in a timely manner. Generally, the higher the education degree means the better the skill, employment, and development opportunities, and the stronger the ability to repay the principal and interest in a timely manner, and thus, the higher borrower's education degree corresponds to the lower default rate. Historical borrowing times represent the success times of the fundraiser in the online loan market in the past, and the more times mean the fundraiser's more successful experience and have a positive feedback on the borrower repaying the principal and interest in a timely manner and thus corresponds to the lower default rate. The higher the fundraiser's income, the better his ability to repay the principal and interest in a timely manner in the future, and thus, the income level has a reverse relationship with the default rate. Housing and car are one kind of capital; on the one hand, owners have had some capital accumulation, and on the other hand, owners' future rigid expenditures on housing and car are lower than those of nonholders, and thus, both housing and car have a significant negative influence on the default rate. Credit level, education, historical borrowing times, income, housing, and car have significant negative influences on default rates which coincide not only on theory requirements but also in actual situations.

In addition, by integrating Tables 2 and 3, seeing from the size and significance of r's coefficient, introducing other variables reduces greatly r's role, and not only does other information have a relatively strong role on the default rate

but also the role of the interest rate on the default rate can be embodied by other information, namely, the role of the interest rate on the default rate is partly through indirect credit level indexes. Default rates which synthesize various information, and which utilize only the interest rate, can be deduced by the models corresponding to Tables 2 and 3. The difference between the two is the default rate, which is not reflected by the interest rate, namely, extra default risk, and it has been excluded in the interest rates' role, and this kind of default rate does not have an interest rate return and can be discriminated and avoided by rational investors.

5.2. Measuring Interest Rate Eliminating Time Value and Default Factor. We have measured different risks behind the same interest rate and obtained the default rate not fully reflected by the interest rate. Next, we apply the term structure theory of interest rate to the online loan interest rate and measure the interest rate that excluded the default risk according to the relationship between interest rates and default risk, and it is the pure yield not corresponding to the default risk, and then, we further test whether online loan investors can discriminate the kind of pure yield and make rational choices.

Interest rates published in the online loan website have different maturities, and interest rates with different maturities are different according to the term structure theory of the interest rate. In order to strengthen the comparability and analyze more specifically the investors' yield discriminating ability, we transfer the interest rate into a continuous compound form according to the formula $[\text{Ln}(1 + t * r_t)]/t$, and thus, heterogeneity effect of different maturities is eliminated.

Since the default rate is ex post variable and cannot be determined completely beforehand, namely, the default rate cannot be measured directly, thus, we apply indirect methods: method 1 only using credit factor and method 2 using various factors. Generally, the default risk is mainly decided by borrowers' credit risk level, and thus, borrowers enact the item's interest rate level according to their own credit risk level, and investors decide whether to accept the interest rate level published by borrowers according to the borrowers' credit level. The Renrendai website publishes a credit level (cl) index according comprehensively to borrowers' various information, and the two indexes are directly

TABLE 3: Multifactor estimated results of heteroskedasticity Probit model of default.

(a)

Variables	dy/dx	Explained variable: whether default
		z-statistics and p value
r	0.0002	1.56 (0.119)
cl	-0.0019	-8.29 (≤0.001)***
lnta	0.0021	5.41 (≤0.001)***
dd	0.0002***	6.61 (≤0.001)***
y	0.0001	2.58 (0.010)***
e	-0.0008	-2.32 (0.020)**
m	0.0002	0.27 (0.785)
hb	-0.0030	-18.22 (≤0.001)***
ho	0.0177	26.60 (≤0.001)***
lni	-0.0014	-3.82 (≤0.001)***
Housing (h)	-0.0016	-2.81 (0.005)***
c	-0.0013	-1.99 (0.047)**
Testing whether existing heteroskedasticity		
Null hypothesis	Lr-statistics and p value	Conclusions
Homoskedasticity	8971.15 (≤0.001)	Rejecting null hypothesis, namely, existing heteroskedasticity

(b)

Variables	Variance equation	
	Coefficient estimated values	z-statistics and p value
r	0.0323	4.1500 (≤0.001)***
cl	0.0326	2.3100 (0.0210)**
lnta	-0.0805	-3.6600 (≤0.001)***
dd	-0.0335	-14.3000 (≤0.001)***
y	-0.0071	-2.5300 (0.0110)**
e	0.0109	0.5000 (0.6180)
m	-0.0046	-0.1100 (0.9130)
hb	0.0413	10.8300 (≤0.001)***
ho	0.5852	30.8900 (≤0.001)***
lni	0.0908	4.0400 (≤0.001)***
h	0.1730	4.4800 (≤0.001)***
c	0.0603	1.4900 (0.1350)

Note: ∗∗∗ means that the corresponding coefficient is significant at 1% significance level, and ∗∗ means that the corresponding coefficient is significant at 5% significance level.

related borrowers' credit level: historical borrowing times (hb) and historical overdue times (ho). Historical borrowing times (hb) reflect the times which borrowers have issued items and succeeded to be full in the Renrendai website. The more historical borrowing times mean borrowers having better historical records and being accepted by online loan investors in the past, and this shows indirectly borrowers having a higher credit level and more probability to repay the principal and interest rate in a timely manner. Historical overdue times (ho) reflect the times borrowers raise funds successfully in Renrendai platform but cannot repay

the principal or interest timely in the end; the more historical overdue times indicate borrowers having the worse records previously, and this indirectly shows that the borrower's credit level is not high, and there is relatively bigger probability to being overdue. To reflecting the impact of the credit risk on the interest rate fully, besides the credit level, historical borrowing times and historical overdue times are also introduced as supplementary variables.

According to Table 4, credit factors have important decisive effects on interest rates, and the credit level and interest rate level are in a significant negative relationship. This is in

Table 4: The influencing results of credit factors on continuous compound interest.

Variables	Explained variable: continuous compound interest	
	Coefficient estimated values	z-statistics and p value
cl	-0.2894	-103.58 (≤0.001)***
hb	-0.0007	-1.29 (0.197)
ho	0.0559	37.78 (≤0.001)***
Constant term	12.7960	795.30 (≤0.001)***

Note: *** means that the corresponding coefficient is significant at 1% significance level.

Table 5: Influencing factors of continuous compound interests.

Variables	Explained variable: continuous compound interest	
	Coefficient estimated values	t statistics and p value
cl	-0.2203	-63.87 (≤0.001)***
lnta	-0.2729	-44.83 (≤0.001)***
dd	0.0022	5.03 (≤0.001)***
y	0.0050	11.37 (≤0.001)***
e	-0.0599	-12.83 (≤0.001)***
m	-0.0265	-3.34 (0.001)***
hb	-0.0088	-15.63 (≤0.001)***
ho	0.0588	40.19 (≤0.001)***
lni	0.0058	1.35 (0.175)
h	-0.0304	-3.91 (≤0.001)***
c	0.1021	12.40 (≤0.001)***
Constant term	15.1065	268.56 (≤0.001)***

Note: *** means that the corresponding coefficient is significant at 1% significance level.

accordance with economic theories, and the higher the credit level indicates the lower the borrower's credit risk and thus corresponds necessarily to the smaller interest rate level. Investors select the interest rate according to not only the credit level but also other credit factors, and the significant positive effect of historical overdue times on the interest rate indicates that maybe the higher historical overdue times correspond to the higher default rate and need the higher yield compensation, and thus items with higher historical overdue times generally correspond to a higher interest rate. This indirectly indicates that market investors are clever and can discriminate credit risk and have the willingness to take part in items with higher risk only when are given higher yield compensation.

The influence of the item credit risk on interest rate is mainly indirectly measured using factors which have direct relationship with credit in Table 4. Although default risks behave mostly as credit risks, other factors also have certain influence on default rate. To measure fully the impact of default risks on interest rates, we further consider more factors and indirectly measure the impact using various factors, and the econometric results are seen in Table 5.

Table 5 reflects the driving effects of borrower's information published by the online loan platform on the item's interest rate, and these factors indirectly embody the borrower's default risk. Testing results in Table 5 are also in accordance with economic theories and actual situations: credit level, education, historical borrowing times, and housing have significant negative effects on interest rates, and this goes along with empirical results of Table 3, and these factors also have significant negative effects on default risks, and all these embody low default risks which correspond to low interest rates. Deadline for repayment, year, and historical overdue times have significant positive effects on the interest rate, and this also goes along with the empirical results of Table 3, and these factors also have significant positive effects on default risks, and all these embody high default risks which correspond to high interest rates.

Term structure theory of the interest rate is applied to decide the interest rate level, and the continuous compound interest is calculated. To further eliminate the credit level's impact, the interest rates, corresponding to a certain default risk, are calculated according to the models corresponding to Tables 4 and 5, respectively. The differences between contin-

uous compound interests and the two former interest rates represent pure yields deducted by default risks and are expressed by $\Delta r1$ and $\Delta r2$, and thus, extra yields after the deduced term structure and default risk factor are obtained.

5.3. Research on Participant Numbers and Time for Being Full. Although online loan interest rates are decided voluntarily by the fundraiser at the government-given range, the investors can also determine the final bargaining interest rate level by the mode of "vote with feet." After the interest rate level is decided by the fundraiser, the investor will make the decision whether to invest comprehensively according to the deadline and risk level, and generally, investors make the decision from the two angles of yield and risk, and they chase the minimal risk at the same yield or the maximal yield at the same risk.

As to the default risk, Liao et al. [7] get the default risk which has no interest rate return by excluding the default risk reflected by the interest rate factor, and thus, the impact of the interest rate can be eliminated when measuring the impact of the default risk on the investor's behavior. But as to the interest rate factor, they have not considered the term structure and default risk level, and thus, the item's interest rate may contain the corresponding term and default risk factors. We first apply the term structure theory of the interest rate in the macroeconomic field to online loan interest rates and get the continuous compound interest to eliminate the influences of different deadlines and then compare it with the continuous compound interest corresponding to the default factors and further get the pure yield deducted the impacts of different maturities and default risk. We measure investors' weighing behaviors, respectively, based on the default risk not reflected by the interest rate and pure yield after deducting the impacts of maturity and default risk.

TABLE 6: Regression results to time for being full.

| Explanatory variables | Explained variable: unit financing's time for being full | | | | |
	Equation (1) Estimated values and p value	Equation (2) Estimated values and p value	Equation (3) Estimated values and p value	Equation (4) Estimated values and p value	Equation (5) Estimated values and p value
Mean equation					
Constant term	0.1037 (≤0.001)***	0.1006 (≤0.001)***	0.1189 (≤0.001)***	0.0864 (≤0.001)***	0.0865 (≤0.001)***
Δr_1	-0.0610 (≤0.001)***			-0.0077 (≤0.001)***	
Δr_2		-0.0856 (≤0.001)***			-0.0106 (≤0.001)***
ΔP_d			1.0305 (≤0.001)***	0.0375 (≤0.001)***	0.0346 (≤0.001)***
ar(1)	0.8505 (≤0.001)***	0.8588 (≤0.001)***		0.9682 (≤0.001)***	0.9688 (≤0.001)***
ma(1)	-0.6445 (≤0.001)***	-0.6537 (≤0.001)***	-0.0014 (≤0.001)***	-0.7945 (≤0.001)***	-0.7952 (≤0.001)***
Variance equation					
Constant term	0.0161 (≤0.001)***	0.0161 (≤0.001)***		0.0001 (≤0.001)***	0.0001 (≤0.001)***
RESID(-1)^2	0.0779 (≤0.001)***	0.0824 (≤0.001***)	0.2176 (≤0.001)***	0.1010 (≤0.001)***	0.1000 (≤0.001)***
GARCH(-1)	0.8912 (≤0.001)***	0.8886 (≤0.001)***	0.7824 (≤0.001)***	0.8974 (≤0.001)***	0.8984 (≤0.001)***
Residual test					
F-statistics of heteroskedasticity test	0.0549 (0.8147)	0.0357 (0.8501)	0.1607 (0.6885)	0.6629 (0.4155)	0.7008 (0.4025)
DW value of autocorrelation test	2.1004	2.0994	1.7299	1.9312	1.9290
Root inverse of ar and ma	Within unit circle	Within unit circle	Within unit circle	Within unit circle	Within unit circle
ADF statistics of stationarity test	-36.7565 (≤0.001)***	-36.3303 (≤0.001)***	-30.7849 (≤0.001)***	-31.2582 (≤0.001)***	-31.2583 (≤0.001)***

Note: ∗∗∗ means that the corresponding coefficient is significant at 1% significance level.

As to the explained variables time and participant number for being full, Liao et al. [7] apply directly the value corresponding to each item and take the raising amount as the explaining variable.

Considering that time and participant number for being full will certainly be affected by the amount being raised, and to measure better the investors' risk and yield discrimination abilities, we use time and participant number corresponding to unit financing, namely, time and participant number of each item divided by the amount being raised by each item. Empirical results are seen in Tables 6 and 7, and they measure, respectively, online loan investors' weighing behaviors between risk and yield through time and participant number for being full.

According to Table 6, first, it is necessary to consider the model heteroskedasticity. On the one hand, coefficients of AR and GARCH terms in the variance equation are significant, and this illustrates that it is necessary to consider heteroskedasticity; on the other hand, residual testing results show that there no longer exists heteroskedasticity in models after considering the first-order GARCH effect, and at the same time, models' residuals are stationary, and all these explain that the models are effective. Second, the default rate not reflected by the interest rate has a significant positive relationship with time for being full, and the bigger default risk corresponds to the longer time for being full, and thus, hypothesis 1 is verified. The pure yield after deducting the effects of different terms and credit risk levels has a signifi-

cant negative relationship with time for being full, and the higher yield corresponds to the shorter time, and thus, hypothesis 2 is verified. The front empirical tests are about the time for being full, but the preference of the investor for the item appears in not only time but also participant number for being full. Next, we research how the default risk not reflected by the interest rate and the pure yield after deducting the effects of different terms and credit risk levels influence the participant number for being full.

According to Table 7, the model heteroskedasticity effect can be measured effectively by the GARCH(1,1) model. On the one hand, the equation coefficients are significant, and this indicates that it is necessary to use the GARCH(1,1) model; on the other hand, it is not necessary to use the higher-order GARCH model by model residual tests, and in addition, model residuals also pass the stationarity test, and all these indicate the models' effectiveness.

Results in Table 7 show that the default rate not reflected by the interest rate has a significant positive relationship with the participant number, and this indicates that the investor can discriminate the default risk not compensated by the yield and make prudent measures and generally reduce the item's investing amount, and thus, more participants are needed to complete the full bid, and hypothesis 3 is fulfilled. The pure yield after deducting the term and credit risk has a significant negative relationship with the participant number for being full, and this indicates that the investor prefers more the item with the higher pure yield, and

Table 7: Regression results to participant number for being full.

Explanatory variables	Explained variable: unit financing's participant number for being full				
	Equation (6) Estimated values and p value	Equation (7) Estimated values and p value	Equation (8) Estimated values and p value	Equation (9) Estimated values and p value	Equation (10) Estimated values and p value
Mean equation					
Constant term	0.0872 ($\leq$0.001)***	0.0873 ($\leq$0.001)***	0.0864 ($\leq$0.001)***	0.0864 ($\leq$0.001)***	0.0865 ($\leq$0.001)***
Δr_1	-0.0096 ($\leq$0.001)***			-0.0077 ($\leq$0.001)***	
Δr_2		-0.0121 ($\leq$0.001)***			-0.0106 ($\leq$0.001)***
ΔP_d			0.0465 ($\leq$0.001)***	0.0375 ($\leq$0.001)***	0.0346 ($\leq$0.001)***
ar(1)	0.9696 ($\leq$0.001)***	0.9700 ($\leq$0.001)***	0.9665 ($\leq$0.001)***	0.9682 ($\leq$0.001)***	0.9688 ($\leq$0.001)***
ma(1)	-0.7920 ($\leq$0.001)***	-0.7928 ($\leq$0.001)***	-0.7919 ($\leq$0.001)***	-0.7945 ($\leq$0.001)***	-0.7952 ($\leq$0.001)***
Variance equation					
Constant term	0.0001 ($\leq$0.001)***	0.0001 ($\leq$0.001)***	0.0001 ($\leq$0.001)***	0.0001 ($\leq$0.001)***	0.0001 ($\leq$0.001)***
RESID(-1)^2	0.0990 ($\leq$0.001)***	0.0981 ($\leq$0.001)***	0.1023 ($\leq$0.001)***	0.1010 ($\leq$0.001)***	0.1000 ($\leq$0.001)***
GARCH(-1)	0.8978 ($\leq$0.001)***	0.8988 ($\leq$0.001)***	0.8957 ($\leq$0.001)***	0.8974 ($\leq$0.001)***	0.8984 ($\leq$0.001)***
Residual test					
F-statistics of heteroskedasticity test	0.7385 (0.3901)	0.7781 (0.3777)	0.6376 (0.4246)	0.6629 (0.4155)	0.7008 (0.4025)
DW value of autocorrelation test	1.9312	1.9289	1.9374	1.9312	1.9290
Root inverse of ar and ma	Within unit circle	Within unit circle	Within unit circle	Within unit circle	Within unit circle
ADF statistics of stationarity test	-31.2087 ($\leq$0.001)***	-31.2277 ($\leq$0.001)***	-31.7400 ($\leq$0.001)***	-31.2582 ($\leq$0.001)***	-31.2583 ($\leq$0.001)***

Note: ∗∗∗ means that the corresponding coefficient is significant at 1% significance level.

thus, less participants are needed to complete the full bid and hypothesis 4 is fulfilled.

The front hypotheses are verified by empirical tests, and the following results can be obtained: First, interest rates and risks mingle with each other, online platform interest rates embody default risks, investors can give trade-offs, and incomplete market-oriented online interest rates reflect the matching result well between the investor and the fundraiser. Second, it is necessary to consider model heteroskedasticity in econometric methods and the blending effect of the interest rate and the default risk in index measurements the impacts of different terms and credit risk levels on interest rates, and the impacts that have been reflected by the interest rate as well as the default rate. Third, the investor's preference for the item embodies directly the time and indirectly the participant number for being full, and generally, if the investor prefers the item more, the shorter the time and the less the participant number needed to be full. Fourth, the investor can discriminate not only default risk not reflected by the interest rate but also pure yield after deducting impacts of different terms and credit risks.

Empirical tests have fulfilled the four hypotheses in the paper's front and verified whether the investor who prefers the item directly embodies the time and specifically shows that the more the preference for the item, the shorter the time for being full, and indirectly embodies the participant number and specifically shows that the more preference for the item, the less the participant number for being full. Com-

pared with the results of Liao et al. [7], both the influencing direction and the significance of different default risks behind the same interest rate on time and participant number for being full are consistent, but either the influencing direction or the significance of the interest rate on time and participant number for being full is inconsistent. It may be due that Liao et al. [7] directly use interest rates without deducting default risks, and thus, the empirical tests show that interest rates have no significant impact on time for being full. On the one hand, our empirical results' credibility is verified; on the other hand, not only should default risks deduct the parts reflected by interest rates but also interest rates should deduct the parts used to compensate the default risk.

5.4. Robustness Test. In the front, we only consider the two factors, the interest rate and the default rate, when researching the time and participant number for being full, and to test robustly, we further consider family income and test adding other variables whether they have some impacts on the conclusions.

According to Table 8, after introducing the income factor, although the coefficient's sizes change slightly, significance and signs of coefficients are the same with the former without considering the income factor, and effectiveness and robustness of our empirical testing results are verified. Additionally, the income level has a significant negative effect on time and participant number for being

Table 8: Regression results to time and participant number for being full.

Explanatory variables	Explained variable: time for being full		Explained variable: participant number for being full	
	Equation (11)	Equation (12)	Equation (13)	Equation (14)
	Estimated values and p value	Estimated values and p value	Estimated values and p value	Estimated values and p value
Mean equation				
Constant term	0.6105 (≤0.001)***	0.4612 (≤0.001)***	0.1266 (≤0.001)***	0.1228 (≤0.001)***
Δr_1	-0.0472 (≤0.001)***		-0.0080 (≤0.001)***	
Δr_2		-0.0361 (≤0.001)***		-0.0104 (≤0.001)***
ΔP_d	1.1429 (≤0.001)***	1.3552 (≤0.001)***	0.0364 (≤0.001)***	0.0340 (≤0.001)***
lni	-0.0538 (≤0.001)***	-0.0401 (≤0.001)***	-0.0043 (≤0.001)***	-0.0039 (≤0.001)***
ar(1)	0.0233 (≤0.001)***	0.9949 (≤0.001)***	0.9686 (≤0.001)***	0.9691 (≤0.001)***
ma(1)		-0.9917 (≤0.001)***	-0.7951 (≤0.001)***	-0.7958 (≤0.001)***
ar(2)	0.8968 (≤0.001)***			
ma(2)	-0.8721 (≤0.001)***			
Variance equation				
Constant term			0.0001 (≤0.001)***	0.0001 (≤0.001)***
RESID(-1)^2	0.0615 (≤0.001)***	0.2222 (≤0.001)***	0.0993 (≤0.001)***	0.0984 (≤0.001)***
GARCH(-1)	-0.0614 (≤0.001)***	-0.2219 (≤0.001)***	0.8983 (≤0.001)***	0.8992 (≤0.001)***
RESID(-2)^2	1.7409 (≤0.001)***	1.4562 (≤0.001)***		
GARCH(-2)	-0.7410 (≤0.001)***	-0.4565 (≤0.001)***		
Residual test				
F-statistics of heteroskedasticity test	0.0718 (0.7888)	0.1073 (0.7432)	0.6906 (0.406)	0.7312 (0.3925)
DW value of autocorrelation test	1.8636	1.8412	1.9292	1.9271
Root inverse of ar and ma	Within unit circle	Within unit circle	Within unit circle	Within unit circle
ADF statistics of stationarity test	-34.3592 (≤0.001)***	-42.9152 (≤0.001)***	-31.3891 (≤0.001)***	-31.4052 (≤0.001)***

Note: *** means that the corresponding coefficient is significant at 1% significance level.

full, and the investor prefers the item corresponding to the higher income level; thus, the full bid needs a relatively shorter time and fewer participants. Our previous basic hypotheses are verified further, namely, investors' preferences for items directly embody the time and express a shorter time needed and indirectly embody the participant number and are expressed as a less participant number needed. Term structure of the interest rate is applied to the online loan interest rate, and the impact of different deadlines is eliminated using continuous compound interest. Further, the annual compound interest is applied (to save space, the testing results are omitted), and the results indicate that signs and significance of coefficients are the same with continuous compound interest except slight changes of coefficient sizes, and effectiveness and robustness of empirical testing results are ensured.

6. Conclusions

Risk levels not reflected by interest rates and pure yields after deducting the term and risk factors are applied to measure investors' discrimination abilities on risks and yields, and online loan investors' discrimination abilities and matching consciousness on risks and yields are researched from the angles of the time and participant number for being full, and the main conclusions and suggestions are obtained as follows.

First, investors have relatively strong matching consciousness, and investors have not blindly selected the item with the high interest rate. Although the higher interest rate corresponds to the higher default risk, the interest rate cannot completely reflect the default risk, and other information are also helpful for reflecting default risk, and investors as passive receivers of online loan rate cannot directly decide the interest rate level, but they can take part in the final interest rate decision by the mode of "vote with feet."

Second, investors' preferences to items embody directly the time and indirectly the participant number for being full. Generally, if the investor prefers the item more, the shorter the time and the less participant number needed for being full. Yield, default rate, and borrower's income are all in accordance with the characteristics.

Third, investors have relatively strong risk discrimination abilities and can identify default risks not reflected by interest rates. If interest rates have been completely market-oriented, then default risks can be reflected entirely by interest rates, and Chinese interest rates are incompletely market-oriented and the same interest rate may correspond to the different default risks, and investors identify this kind of default risk using other information besides interest rates. Investors' discrimination consciousness on this kind of default risk embody directly the time and indirectly the participant number, and items with the bigger kind of default risk need longer time and more participant number for being full.

Fourth, investors have relatively strong yield discrimination abilities and can identify pure yields after deducting deadline and credit risk factors. Investors' discrimination on pure yields embody directly the time and indirectly the participant number, and items with the bigger pure yields need the shorter time and the less participant number for being full.

Fifth, the online loan platform has a relatively strong self-purification function, and historical information (borrowing times and overdue times) have significant forecasting effect on the default risk. The self-purification function of the online loan platform guarantees its sustainable development, but there is further subdivided space in credit valuing, and it is necessary to strengthen the discrimination function of credit valuing to the default risk.

In short, although the online loan interest rate is incompletely market-oriented, it embodies demands of borrowers and lenders and is the weighing result of borrowers and lenders. Investors can identify the different default risks behind the same interest rate and the different yield level behind the same credit risk. Our empirical results show the following: the completely market-oriented online loan interest rate has both actual basis and necessity. Investors having relatively strong rational consciousness can identify not only the different credit risk behind the same interest rate but also the pure yield after deducting the different deadline and credit risk's effects, and these offer feasible actual basis for marketization of online loan interest rates. Additionally, when the default risk has not been compensated by the interest rate or the pure yield, after deducting deadline and credit risk effects, is too low, the item being full bid needs more participants and a longer time. If investors' number is limited in the online loan market, many items may be un-full bid, and the effective funds needed may be suppressed. Complete market-oriented interest rates can make the item be full bid by use of shortening time and reducing participant number for being full, thus satisfying better borrowers' financing needs. ,Next interest rate marketization reform should be strengthened further, and interest rate marketization and competing mechanism should be loosened and introduced in online loan markets, and thus, the price decision mechanism of the interest rate is played better, and investors' behavior is guided better by the interest rate.

In addition, the online loan platform as a shared information platform can exert the roles of information intermediary, transaction cost reduction, and fund allocation effectiveness improvement, but compared with the traditional financial market, the network information asymmetry is serious. And thus, in the next, information disclosure especially borrowers' historical information should be perfected and strengthened further, and online loan investors' information collection ability and yield and risk identification abilities using existed information should be enhanced.

The paper has researched investors' recognition abilities on default and yield using big financial data but the mathematics deduction has not been offered, and in the future, we will give a rigorous mathematical proof.

Data Availability

The raw data can be obtained through the crawler method after registering a personal account in the Renrendai platform. And the data are also available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Klafft, "Peer to peer lending: auctioning microcredits over the internet," in *Proceedings of the international conference on information systems, technology and management*, IMT, Dubai, 2008 https://ssrn.com/abstract=1352383.

[2] Y. Li, Y. Guo, and W. Zhang, "Analysis on the influencing factors of lending success rate in China's P2P micro finance market," *Journal of Financial Research*, vol. 7, pp. 126–138, 2013.

[3] Y. Li, Y. Gao, Z. Li, Z. Cai, B. Wang, and Y. Yang, "The influence of borrower's description on investors' decision-analyze based on P2P online lending," *Economic Research Journal*, vol. S1, pp. 143–155, 2014.

[4] D. Liu, D. J. Brass, Y. Lu, and D. Chen, "Friendship in online peer-to-peer lending: pipes, prisms, and relational herding," *MIS Quarterly*, vol. 39, no. 3, pp. 729–742, 2015.

[5] Q. Wan, D. Chen, and W. Shi, "Online peer-to-peer lending decision making: model development and testing," *Social Behavior and Personality*, vol. 44, no. 1, pp. 117–130, 2016.

[6] R. Iyer, A. I. Khwaja, E. F. P. Luttmer, and K. Shue, "Screening in new credit markets: can individual lenders infer borrower creditworthiness in peer-to-peer lending?," 2009, https://ssrn.com/abstract=1570115.

[7] L. Liao, M. Li, and Z. Wang, "The intelligent investor: not-fully-marketized interest rate and risk identify," *Economic Research Journal*, vol. 2014, no. 7, pp. 125–137, 2014.

[8] C. Serrano-Cinca, B. Gutiérrez-Nieto, and L. López-Palacios, "Determinants of default in P2P lending," *PloS One*, vol. 10, no. 10, article e0139427, 2015.

[9] R. Emekter, Y. Tu, B. Jirasakuldech, and M. Lu, "Evaluating credit risk and loan performance in online peer-to-peer(P2P) lending," *Applied Economics*, vol. 47, no. 1, pp. 54–70, 2015.

[10] R. Ge, J. Feng, B. Gu, and P. Zhang, "Predicting and deterring default with social media information in peer-to-peer lending," *Journal of Management Information Systems*, vol. 34, no. 2, pp. 401–424, 2017.

[11] Y. Liu, Q. Zhou, X. Zhao, and Y. Wang, "Can listing information indicate borrower credit risk in online peer-to-peer lending?," *Emerging Markets Finance & Trade*, vol. 54, no. 13, pp. 2982–2994, 2018.

[12] E. Ravina, "Beauty, personal characteristics and trust in credit markets," 2007, https://ssrn.com/abstract=972801.

[13] S. Freedman and G. Z. Jin, "Do social networks solve information problems for peer-to-peer lending? Evidence from Prosper.com," in *Working Paper #08-43*, University of Maryland & NBER, 2008.

[14] F. Guo, "Network nickname, P2P lending success rate and default rate," *Economic Science*, vol. 6, pp. 102–116, 2016.

[15] Z. Yue, Q. Zhou, and X. Yang, "Crowdfunding, information screening and market efficiency——an empirical study based on Renrendai," *Economic Perspectives*, vol. 1, pp. 54–62, 2016.

[16] F. Guo, "Marriage and credit identification: evidence from P2P lending platform," *Modern Economic Science*, vol. 39, no. 3, pp. 40–47, 2017.

[17] H. Zhang and H. Cai, "The value of tittle-evidence from P2P lending," *China Economic Quarterly*, vol. 17, no. 4, pp. 1525–1556, 2018.

[18] J. J. Xu and M. Chau, "Cheap talk? The impact of lender-borrower communication on peer-to-peer lending outcomes," *Journal of Management Information Systems*, vol. 35, no. 1, pp. 53–85, 2018.

[19] F. Caldieraro, J. Z. Zhang, M. Cunha Jr., and J. D. Shulman, "Strategic information transmission in peer-to-peer lending markets," *Journal of Marketing*, vol. 82, no. 2, pp. 42–63, 2018.

[20] J. Hu, J. Li, and B. Zhang, "Does online peer-to-peer lending help to achieve inclusive finance?," *The Journal of World Economy.*, vol. 11, pp. 169–192, 2018.

[21] G. Babaei and S. Bamdad, "A neural-network-based decision-making model in the peer-to-peer lending market," *Intelligent Systems in Accounting, Finance & Management*, vol. 27, no. 3, pp. 142–150, 2020.

[22] S. M. Freedman and G. Z. Jin, *Learning by Doing with Asymmetric Information: Evidence from Prosper.com*, NBER Working Paper No. 16855, 2011.

[23] M. Gao, J. Jiang, J. Chen, and Y. Liu, "Gender effect, overconfidence and P2P investor behavior," *Journal of Financial Research*, vol. 11, pp. 96–111, 2017.

[24] G. Dorfleitner, E.-M. Oswald, and R. Zhang, "From credit risk to social impact: on the funding determinants in interest-free peer-to-peer lending," *Journal of Business Ethics*, vol. 170, no. 2, pp. 375–400, 2021.

[25] J. Hu and W. Song, "The rational sense and balanced behavior of P2P investors: empirical analysis based on transaction data from Renrendai.com," *Journal of Financial Research*, vol. 7, pp. 86–104, 2017.

[26] G. A. Akerlof, "The market for "lemons": quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, vol. 84, no. 3, pp. 488–500, 1970.

[27] W. Bagehot, "The only game in town," *Financial Analysts Journal*, vol. 27, no. 2, pp. 12–14, 1971.

[28] D. Kahneman, *Attention and Effort*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1973.

[29] L. Peng and W. Xiong, "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, vol. 80, no. 3, pp. 563–602, 2006.

[30] A. C. Harvey, "Estimating regression models with multiplicative heteroscedasticity," *Econometrica*, vol. 44, no. 3, pp. 461–465, 1976.