The Existence of Herding in P2P Lending and the Rationality of It — An Empirical Study from China

Xiangyu Wang and Yan Tu

***Abstract*—Online peer-to-peer (P2P) lending is an innovative way of lending, which enables borrowers fetch money from individual lenders online rather than through traditional financial institutions. Using data from Prosper, which is the largest P2P loan platform in the United States, many researchers have found that there is a herding behavior in P2P lending and as opinions vary, the rationality of it is uncertain. We study herding behavior in P2P lending in China. Data is based on PPDai, which is China's largest P2P lending platform. To elaborate the effect of herding, we define a herding index which describes the extent of herding rather than simple bidding numbers as many researchers do. Based on some limitation of current research, we extend our model and empirically test the existence of herding and the rationality of it. The research enriches the study of herding in P2P lending. We divide herding behavior into two aspects, which is beneficial to deep understand the herding and elaborately express the extent of herding, which benefit the examination of the rationality of it.**

***Index Terms*—P2P lending, herding, PPDai, herding index, rationality.**

I. Introduction

The past decade has witnessed a rapid growth of P2P lending, which is accelerated by Web 2.0. In P2P lending, borrowers and lenders transact on online platforms, such as Prosper, Lending Club, PPDai, and Zopa. According to the report of Price water house Coopers in 2015, the US P2P lending platforms issued approximately $5.5 billion in loans in 2014. In China, the volume of P2P loan is even larger and reached $41.3 billion in 2014. P2P lending makes it possible that both borrowers and lenders can transact online without intermediaries such as banks [1]. P2P lending can not only finance SMEs but also finance individuals. P2P lending is beneficial to both borrowers and lenders. Due to financial decentralization, borrowers can get loans directly from lenders and pay low interest rates [2]. In addition, P2P lending can mitigate financial exclusion, i.e., P2P lending can finance borrowers with low credit grades [3]. For lenders, they can obtain more opportunities of investment and get high returns [4].

A typical process of P2P lending is as follows. First, borrowers apply for a loan request called listing, containing the information of loan amount, interest rate, duration and purpose of loan. The website will show the personal information of borrowers including gender, occupation and credit rate. Lenders who intend to lend then bid on the listing. Other choices of bidders will be shown in the page of listing, which is a potential cause of herding.

We intend to solve two problems: first, whether there is a herding behavior in P2P lending platform and second, whether such a herding behavior is rational. Different from current research, we analyze the representation of extent of herding and incorporate an index into the model examining the rationality of herding. We believe the index can perfectly represent the extent of herding.

The rest of the paper is organized as follows. First, we review the literature about P2P lending and herding behavior. Second, we elaborate the limitation of current research and propose our method to improve current work. Last, we describe our data, analyze results and draw a conclusion.

II. Literature Review

A. P2P Lending

P2P lending has attracted both practical and theoretical attention due to its popularity. Research concerning P2P lending can be divided into three parts. First, early research focuses on the reasons for the emergence of online P2P lending. Second, current research examines the determinants of funding success and default, which considered as the basis of decision-making of borrowers and lenders. Third, some research investigates the performance of online P2P loan taking risk into consideration. Above them, the second part is very important and there is extant research about it. Borrowers wish to get loans in low interest rates while lenders make efforts to discern whether borrowers pay loans in time. Research based on the perspective of borrowers is scarce. Borrowers face a tradeoff between loan amount and interest rate, i.e., borrowers may have to set high interest rates high to ensure the funding success [5]. Research based on perspective of borrowers usually use auction theory and provide advice of loan amount and interest rate setting [5][6].

Research based on perspective of lenders usually focus on the factors of decision-making. One of problems in P2P lending is information asymmetry, i.e., lenders don't know the credibility of borrowers for certain as borrowers do [7]. Information asymmetry leads to high risk for lenders. In traditional banking systems, information asymmetry can be lessened due to financial intermediaries, the absence of which makes it acute in P2P lending. Further, the inherent anonymity of online environment intensifies the problem. Since information asymmetry between borrowers and investors leads to moral hazard [8] and adverse selection [9], lenders must be cautious when making decisions.

In P2P lending, lenders use both hard information and soft information about borrowers to make investment

Manuscript received August 25, 2016; revised November 1, 2016.

The authors are with the School of Information , Central University of Finance and Economics, Beijing, China (e-mail: tcxdd@163.com, tuyancufe@163.com).

doi: 10.18178/joebm.2016.4.11.468

decisions [10], [11], [12]. Hard information refers to quantitative information that can be accurately expressed, e.g., demographic information, debt to income ratio, FICO score, credit grade [13]. Hard information is often difficult to obtain, insufficient, or unreliable, so lenders may tend to soft information, which is available in P2P lending platforms and can be diagnostic [14]. Soft information refers to non-standard qualitative information [11], e.g., narrative, appearance, social networking (e.g., virtual community and friendship). Lin et al. (2013) discover that strong social ties and friendship can increase the probability of funding success and decrease the probability of default. The result is consistent with the research of Freedman and Jin (2014). Liu D et al. (2013) further distinguish different effects of three types of friendship, i.e., the pipe effect, the prism effect and relational herding effect.

B. Herding Behavior

We firstly review the research about herding not confined to the context of P2P lending. Herding is a phenomenon usually happens in animal kingdom and the definition of herding behavior is that individuals are strongly influenced by the decisions of others [15] and many researchers find the existence of herding behavior in various contexts, e.g., investment (Graham 1999), online auctions [16], IT adoption [17].

The most prevalent reason why we behave like herd is "information cascades" [18], which is based on the interpretation of conformity preference. "Information Cascades" means private information is too weak to resist the information of mainstream. It occurs especially when individuals can't obtain the full information, which is the context of P2P lending [18].

Many researchers study P2P lending by incorporating herding model into the context of P2P lending. Herd behavior in P2P lending refers to the tendency of lenders to gravitate toward and bid on auction listings with existing bids [16]. The main question can be divided into two dimensions, i.e., whether there exists herding behavior and whether it is rational.

1) The existence of herding in P2P lending

Many researches use data from different P2P lending platforms and testify that there exists herding behavior. For example, Herzenstein M et al. (2011) used data from Prosper and conducted an elaborately study. They found that before loans receive enough bids, a 1% increase in the number of bids increases the likelihood of an additional bid by 15% but after loans receive enough bids and comes to phase of auction, a 1% increase in bids increases the likelihood of an additional bid by only 5%. This diminishing effect of herding is also found by Lee E, Lee B [1], who used data from Popfunding and found that as the level of participation increase, the newer biddings increase tardily.

Apart from empirical study based on various P2P lending platforms, some researchers theoretically testify the existence of herding in P2P lending. Using decision tree model, Luo B, Lin Z.(2013) explained the formation of herding in P2P lending and further found empirical evidence of herding by using data from Prosper.

2) The rationality of herding

When it comes to the research about rationality of herding, conclusions are varied. Herzenstein M et al.(2011) found that though there is adverse effect of herding for buyers in eBay, herding is advantageous for lenders in P2P lending. Dramatically, Luo B, Lin Z.(2013) used data from the same P2P lending platform, i.e., Prosper, and believed that herding will impair benefits of lenders. Other P2P lending platforms are also concerned as for the rationality of herding. Chen D, Lin Z.(2014) used PPDai and found that herding behavior lowers the final interest rate and even worse, increases default rates. They believed that cultural and economic factors play a role in the difference of their results from others.

III. Limitation of Current Research and Model Building

We believe that there are some limitations in the current research which interests us.

A. The Expression of Extent of Herding

To examine the rationality of herding, the expression of extent of herding is essential. Most researchers use numbers of bidders or biddings to express the extent of herding, following the study of online auction websites, e.g., eBay. We don't believe it's appropriate for the context of P2P lending.

First, in online auction websites, as long as there is someone bidding, the auction continues while in P2P lending platforms once the total bidding amount reaches the request amount, the bidding ceases. Such restrict of amount leads to the bias: the number of bidders or biddings can't precisely denote the extent of herding and is relevant to the former amount of bidding. For example, the amount of total bidding has reached 99%, completed by 90 bidders. There are 100 bidders aspiring to bid the left because of herding. Due to restrict of amount, only 1 bidder bid while 99 bidders left can't. The real number of bidders denoting the herding is (90+100) while the observed value is (90+1), 99 bidders are overlooked. There is no bias for online auction websites because we assume that those 100 bidders bid because of herding and bid whatever the final price is and they can all bid successfully.

Second, it's worth discussing the mechanism of end of a listing. In some P2P lending platforms, the listing is closed once it receives enough bids. We call such platforms type A platform. In other platforms e.g., Prosper, the listing is not closed immediately when it receives enough bids. They provide an auction mode and continue the listing(we call it "auction phase"). Additional bids can lower interest rates, which benefit borrowers. We call such platforms type B platform. Apparently those who believe the numbers of bidders or biddings can represent the extent of herding can only conduct the study on type B platforms, which impair the generalization of their results.

Third, even in type B platforms, the number of bidders or biddings can't precisely represent the extent of herding. The number of bidders or biddings before the auction phase and after it holds different effect to the representation of herding. For example, there are 2 listings named listing1 and listing2. The total numbers of bidders are both 100 but in the listing1 there are 90 bidders bid during auction phase(it means that before the auction phase 10 bidders bid the request amount) and 10 in the listing2. Apparently the extent of herding in listing1 is higher than listing2, in which everyone is reluctant to bid more and latently increase the number of bidders.

In sum, the different mechanism between auction and lending means the method used in auction is not quietly appropriate for lending.

B. The Definition of Rationality

We believe in the current research the definition of rationality is ambiguous and not intact. Some researches simply use the variance between final interest rates and initial interest rates to evaluate the rationality. However, such variance is obvious and other factors must be detected. Others consider the default rates but they overlook that bidding on the listing that many others already bid can ensure the success of the listing, which is important because failure of listing impair benefit of bidders for that their illiquidity asset is locked in the platform and lose time value of money. In short, synthetic factors need to be evaluated.

In this paper, we intend to solve two questions: the existence of herding behavior and the rationality of it. The reason for herding in P2P lending can be explained in two aspects. First, the problem of information asymmetry is acute in P2P lending compared to traditional banks and lenders can't deduce the credit status of borrowers alone. They use information of other lenders to form their judge towards the borrowers. Second, the mechanism of P2P lending causes herding because when the process of a listing nearly reaches to the request amount, lending to the listing can assure the listing is successful, which also benefits lenders because their money can be effective as soon as possible. Thus we have hypothesis H1.

H1: There exists herding behavior in P2P lending

We believe that herding behavior of lenders in P2P lending can be reflected in two aspects: the decision of choosing borrowers and the decision of choosing listings.

When lenders choose borrowers, they not only exert their own information towards borrowers but also refer to judgments of others, which can practically be realized by the function of viewing history borrowing records of borrowers in PPDai platform. Successful funding records in the past indicate the recognition of other lenders towards borrowers and a potential lender might leverage such information and make decision. Thus we assume that when times of successful funding records in the past increase, the probability of successful funding increases, which indicate the herding behavior of lenders.

H1a. When times of successful funding records of borrowers in the past increase, the probability of successful funding increases.

When lenders choose listings, they can obtain bidding information from other borrowers in the webpage of the listing. To elaborately prove the existence of herding, we intend to follow Lee E, Lee B. (2012)'s research and use one day as a time span. The dependent variable is the ratio of bidding amount to the request amount in the day for certain listing and the main independent variable is the ratio of accumulated bidding amount to the request amount before the day. Other control variables, e.g., credit rate, interest rate, duration are concerned. For example, the request amount of listing i is \$10000. Before the day j there have been \$5000 bided in the listing and in the day j there are \$3000 bided in it. Thus the value of the independent variable and the dependent variable is 50% and 30% respectively in the observation. If there exists herding behavior, the bidding amount in a day increases when the amount before the day increases. Based on this, we build hypothesis H1b.

H1b. The bidding amount in a day increases when the amount before the day increases

As aforesaid the conclusion of rationality is not the same in different study. Even using the same dataset - Prosper, Herzenstein M et al.(2011) and Luo B, Lin Z.(2013) draw different conclusions. Thus the rationality of herding is debatable and the study of it is necessary. We follow the conclusion of Chen D, Lin Z.(2014) because they use data from the same P2P lending platform and we believe that due to cultural and economic factors herding is irrational.

H2: The herding behavior in P2P lending is irrational

Following the examination of existence of herding, we analyze the rationality of herding in two aspects: herding behavior towards choosing borrowers and herding behavior towards choosing listings.

To examine the rationality of herding we need to measure the extent of herding and examine the relation of default rate and it. Thus we define an herding index which we think is appropriate for representation of the extent of herding. Fig. 2 shows the source of the index of certain listing. The y axis means the ratio of accumulated bidding amount to the request amount and the x axis means the ratio of elapsed time from beginning of the listing to the duration of bidding. The path from O to B represents the process the listing ends. Straight line OB is the most common situation in which the speed of bidding is average while arc OAB represent the herding because in the beginning of the listing the speed is low but when it comes to the end of listing the speed is high. The herding index is the value of the area of the black part, i.e., the area of the part surrounded by Straight line OB and arc OAB. If the extent of herding is acute, the degree of crook of OAB is large and thus the area of black part is large.

Some researchers use average time interval of bidding or relative time elapsed (the proportion of time that had elapsed in the auction when the loan received full funding), which we believe is inappropriate. The reason can be illustrated in Fig. 2 and Fig. 3. Suppose Fig. 2 shows the bidding process of listing A and Fig. 3 shows the bidding process of listing B. We suppose the average speed of bidding is equal between A and B and the only difference is that the speed is first low and then fast in A while first fast then low in B. In the listing B, the bidding speed at first is high but when it comes to the end, the speed is low, which means bidders at first are interested in the listing because the listing itself is attractive and they are afraid not to bid before it ends. The speed of subsequent bidders' biding is lower than normal, which means they are not as in need as former bidders and the former bidding is not related to theirs. In other words, they are not imitating and there is totally no herding behavior. The average time interval of bidding is equal between listing A and listing B but apparently the extent of herding is not. Yet the herding index can distinguish the variance between the listing A and B because the herding index of B, the listing in which herding totally not exist, is negative.

Such graphic and method of calculation is similar to Gini coefficient. Gini coefficient measures the extent of inequality of income, the calculation of which is from Lorenz curve. In a Lorenz curve, the x axis denotes the percentage of population from the poorest to the richest and the y axis denotes the percentage of welfare. Suppose figure 2 represent a Lorenz curve and straight line OB means absolute equality of income and arc OAB means actual distribution of income. The Gini coefficient is the result of area of black part divide 0.5.

We incorporate Gini coefficient and construct a herding index because we believe the essence of herding is the inequality of bidding speed while Gini coefficient exactly measures the extent of inequality of distribution. However there is misuse concern, i.e., in Lorenz curve the x axis denote the percentage of population that is already ranked from the poorest to richest and arc OAB will always be under straight line OB while in herding index arc OAB can surpass straight line OB. We believe that there is some tiny difference between Gini coefficient and herding index. Gini coefficient focuses on the inequality of distribution but herding not only means the inequality of bidding speed but also the concept that the bidding speed is slow in the beginning but fast in the end. If the speed is fast at first but slow in the end herding is not significant while the inequality of bidding speed is. When arc OAB surpass straight line OB there is punishment for representation of herding.

There may not be matching perfect arc OAB in the real transaction context. To calculate the herding index, we can use calculus method, i.e., we divide the black part in figure2 into some echelons and calculate the sum of the area of echelons. Figure 3-5 show three examples of bidding process in the real transaction, indicating low herding, normal herding and high herding respectively. The ids of listing are 404404,404406 and 428701. The information of listings can be viewed in website. URL can be inferred based on id of the listing (e.g., the information of listing 404404 can be viewed on http://www.ppdai.com/list/404404). After calculation, we find that the corresponding herding index is -0.2995, -0.0074 and 0.3323 respectively.

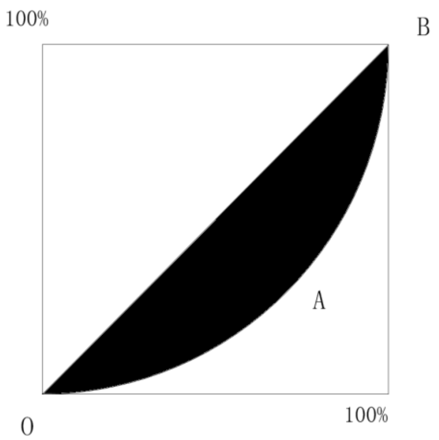

Fig. 1. The bidding process of listing A

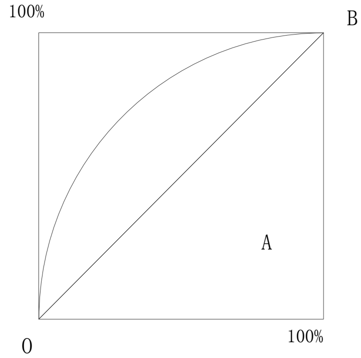

Fig. 2. The bidding process of listing B.

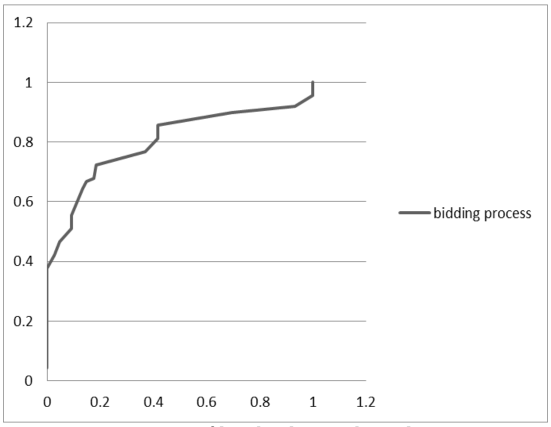

Fig. 3. Listing of low herding in the real transaction.

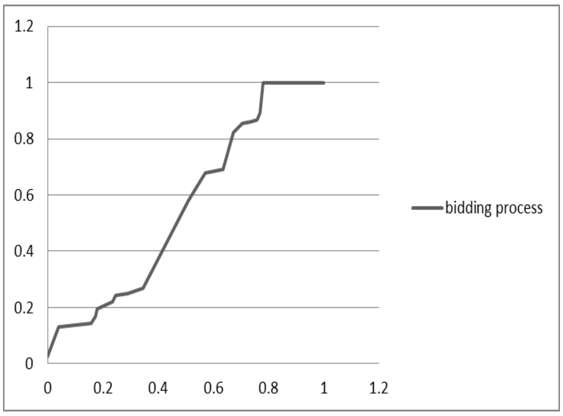

Fig. 4. Listing of normal herding in the real transaction.

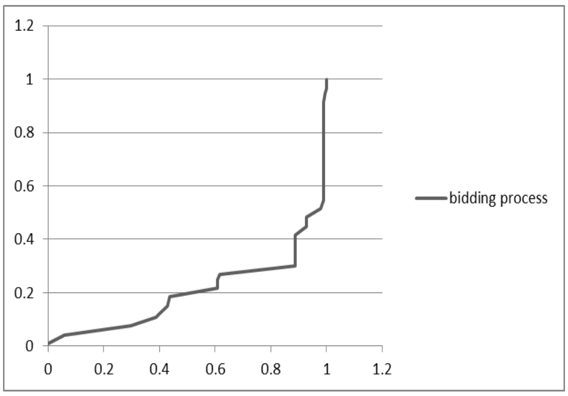

Fig. 5. Listing of high herding in the real transaction.

IV. DATA

Data is from PPDai, the largest P2P lending platform. We captured the data from 2014.1.1 to 2014.3.31 and all transaction records during this time interval are recorded in the database. The mechanism of URL in PPDai brings out great convenience because all listing IDs are parts of URL of the listing. We first find out the range of loan ID because loan ID is determined according to time order. Then we use python to record all the web text from those pages. To analyze these texts and derive the data needed, we write program to handle it. The main data item includes type, title, the percent of progress, loan amount, interest rate, duration, bidding list, purpose. The web page of listing does not consist of result of repayment, i.e., whether there is default or not, thus we captured the data about borrowers concerned. The same procedure is conducted. The main data item includes age, gender, occupation, credit rate, authentication status. There are 102430 listings and 79709 borrowers incorporated in the dataset. The number of bidding record is 565539. Among all listings there are 15906 (15.5%) listings successfully funded. The variables and simple descriptive statistic results are shown in Table II and Table II.

TABLE I: QUANTITATIVE VARIABLES

Type	Variable	Max	Min	Average	STD
listing	loanAmount(RMB)	400000	1000	3956	9426
	Annual interest rate	24	8	15.7	3.9
	Duration(month)	12	1	9	2.8
	# of Bidders	1516	0	5	23
Borrower	# of Successful borrowing	728	0	11	66
	# of failed borrowing	31	0	2.45	2.36
	# of bidding	40918	0	23	473
	Weighted lending rate	25	0	1.41	4.66

TABLE II: QUALITATIVE VARIABLES

Type	Number	Variable	Dummy Variable	Number
listing	102430	Type	Type A	15643
			Type B	5229
			Type C	5564
			Type D	751
Borrower	79709	Credit rate	AAA	690
			AA	13523
			A	57
			B	488
			C	1146
			D	5115
			E	47441
			F	11249
		gender	Male	68855
			Female	10854
		occupation	student	2684
			wage-earners	45532
			Private business owners	17527
			Online shop owners	3836
			Others	10130

Table I shows the quantitative variables while Table II shows the qualitative variables. Some information can be discovered. First, the largest request amount of listing is 400000 RMB, which is a considerable amount but the average amount is 3956, which means the request amount is not large in most of listings. Second the highest interest rate is 24%, and the average of it is 15.7%. The rate in P2P lending is much higher than the rate of checking accounts, which is only nearly 0.35% at that time. Such variance of rates attracts many lenders to P2P lending platforms. Third, we discuss the listing type in PPDai. In PPdai, there are 4 types of listing. Type A means platforms compensate for lenders when the listing is overdue. Type B means the money borrowing in the listing can't be withdrawn and can only be used within the platform, e.g., bidding or repayment. Type C means the listing that the payback source is guaranteed, e.g., accounts receivable or money saved in the platform. Type D means listing with guarantee. Forth, we can tell that the main borrowers of P2P lending are male. The reason may be that males have much burden in life than females and need more money. Last we notice that more than half borrowers are wage-earners because the wage they earn may not support their extra need while the occupation is stable for them and make it credible for them to pay back.

V. RESULT

A. The Existence of Herding

We analyze the existence of herding by using logistic model. The main variables are shown in Table III.

1) The existence of herding when lenders choose borrowers

We create logistic regression model in which dependent variable is whether listing is successfully funded and the main independent variable is times of successful borrowing records in the past. Table IV shows the result of the model.

TABLE IV: MAIN VARIABLES

Variable	Meaning
GABC	If the credit grade of borrower is A,B or C and the value is 1, else 0. We set this variable because borrowers with credit grade D,E and F is very few.
SUCCESS	The amout of successful borrowing records in the past
FAILURE	The amout of failed borrowing records in the past
TYPE	Four types of listing
LNAMOUNT	The log of amount. We use logarithm process because the amount is very large
RATE	Interest rate
AGE	The age of borrowers
OCCUPATION	The occupation of borrowers
CER	The certification of borrowers

TABLE V: THE EXISTENCE OF HERDING WHEN LENDERS CHOOSE BORROWERS

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-18.59984	0.628439	-29.59690	0.0000
GABC	5.149296	0.053042	97.07980	0.0000
SUCCESS	0.018199	0.001996	9.115842	0.0000
FAILURE	-0.204898	0.004889	-41.90597	0.0000
TYPEB	0.756571	0.097528	7.757447	0.0000
TYPEC	0.312124	0.089595	3.483729	0.0005
TYPED	2.765786	0.108327	25.53182	0.0000
LNAMOUNT	1.134480	0.050280	22.56341	0.0000
RATE	0.395779	0.003127	126.5839	0.0000
AGE	0.001056	0.000476	2.218445	0.0265
OCCUPATIN_STUDENT	0.175779	0.059845	2.937235	0.0033
OCCUPATION_WAGE_EARNERS	0.112231	0.033785	3.321912	0.0009
OCCUPATION_OSOWNERS	0.268957	0.057165	4.704937	0.0000
OCCUPATION_PRIVATEOWNERS	0.307766	0.038843	7.923381	0.0000

CER_IDENTITY	1.071651	0.475227	2.255028	0.0241
CER_VIDEO	0.483560	0.031295	15.45163	0.0000
CER_EDUCATION	0.248571	0.036469	6.815975	0.0000
CER_MOBILE	0.668838	0.025835	25.88836	0.0000
CER_ONLINEBANKING	1.062544	0.027964	37.99642	0.0000

McFadden R-squared	0.550145	Mean dependent var	0.416195
S.D. dependent var	0.492929	S.E. of regression	0.306796
Akaike info criterion	0.611309	Sum squared resid	9534.632
Schwarz criterion	0.613096	Log likelihood	-30949.31
Hannan-Quinn criter.	0.611851	Restr. log likelihood	-68798.36
LR statistic	75698.10	Avg. log likelihood	-0.305467
Prob(LR statistic)	0.000000		

Obs with Dep=0	59150	Total obs	101318
Obs with Dep=1	42168		

The result shows that the coefficient of successful funding times in the past is significantly positive, which confirms hypothesis 1a.

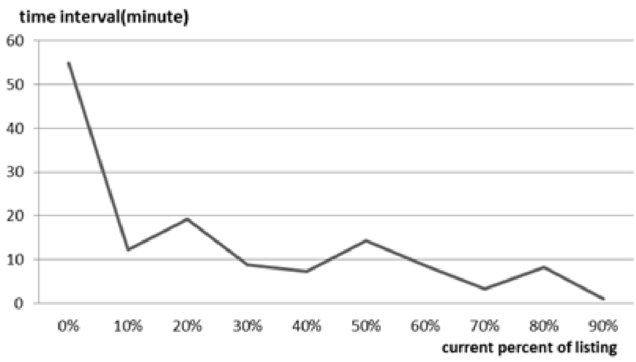

Fig. 6. Evidence of herding.

Other conclusions can be found. First, when the credit grade is high (GABC), the probability of successfully funding is high and the coefficient value is relatively large, which means the credit grade is main concern of lenders when choosing borrowers. Second, when listing is specially certificated, the probability of successfully funding increases. The coefficient of type D is relatively high, because type D means the listing is guaranteed, which is a direct signal of low risk while other types are indirect for lenders. Third, the coefficient of rate is positive, which means lenders pursue high profit. When it comes to age of borrowers, we find that lenders prefer old borrowers and one possible explanation is that old borrowers are more credible. The finding is the same as previous research (Gonzalez and Komarova Loureiro, 2014). Last, the private business owners are more likely to get loans because the coefficient is relatively high. When borrowers are authenticated by PPDai, the probability of successfully funding is high especially when identities of borrowers are authenticated.

2) The existence of herding when lenders choose listings

We draw the figure 6 using our dataset in which x axis means the percent of complement of listing, i.e., the ratio of accumulated bidding to the request amount and y axis means the time interval between two successive bidding, which is a mean value of all data from our dataset. As shown in the figure, when the current percent of complement of listing is less than 10%, the speed of listing is very low and average time interval between two successive bidding is nearly 1 hour. When the bidding proceeds and the percent of complement of listing is 90%, the speed of bidding is very fast and average time interval between two successive bidding is nearly 1 minute. The time interval between two successive bidding decreases when accumulated amount of bidding increases, proving the existence of herding.

In addition, when current percent is low, the decrease of time interval is not acute while when current percent is high the decrease of time interval is inconspicuous, proving the marginal decrease effect of herding. The result is the same as Herzenstein M *et al*. (2011).

TABLE VI: THE RATIONALITY OF HERDING

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.302860	0.029361	10.31501	0.0000
GABC	-0.137964	0.013815	-9.986190	0.0000
SUCCESS	0.000343	2.42E-05	14.16693	0.0000
FAILURE	-0.010680	0.001519	-7.033386	0.0000
TYPEA	0.134429	0.011308	11.88790	0.0000
TYPEB	-0.102351	0.013861	-7.384078	0.0000
TYPED	0.051061	0.020881	2.445383	0.0145
RATE	0.002478	0.001229	2.015790	0.0438
DURATION	0.011603	0.001067	10.87373	0.0000
MALE	0.039104	0.007969	4.907102	0.0000
OCCUPATIN_STUDENT	-0.206123	0.018751	-10.99244	0.0000
OCCUPATION_WAGE_EARNERS	0.043478	0.006031	7.208884	0.0000
CER_EDUCATION	-0.080725	0.007109	-11.35561	0.0000
CER_MOBILE	-0.031996	0.006513	-4.912334	0.0000
TOTALLY_PAYBACK	-0.012228	0.000292	-41.91884	0.0000
DEFAULTBUTPAYBACK	-0.011569	0.001515	-7.635030	0.0000

R-squared	0.275382	Mean dependent var	0.182037
Adjusted R-squared	0.274590	S.D. dependent var	0.385889
S.E. of regression	0.328665	Akaike info criterion	0.613610
Sum squared resid	1482.369	Schwarz criterion	0.622376
Log likelihood	-4199.191	Hannan-Quinn criter.	0.616531
F-statistic	347.6843	Durbin-Watson stat	1.977418
Prob(F-statistic)	0.000000		

B. The Rationality of Herding

We conduct logistic regression model and analyze two aspects both. The dependent variable is default. The result is shown in Table V.

1) Rationality of herding when choosing borrowers

The coefficient of SUCCESS is significantly positive while FAILURE is significantly negative, which is contrary to common sense. We believe one possible explanation is that borrowers with more successful borrowing records can obtain loans easily (see Table 4) and they may use such advantage and make a fortune before leaving the market. Borrowers with more failed borrowing records should have little chance to get loans successfully but once they get loans, they cherish such chance and try to refund to improve their credit status. Thus we believe that the herding behavior of lenders when choosing borrowers is irrational.

Other conclusions can be found. First, when credit grade is low, borrowers tend to default. Second, the coefficient of TYPE A and TYPE D is positive. One possible explanation is as follows. For type A listings, platforms compensate for lenders when the listing is overdue. When borrowers default, platforms will replace borrowers to compensate, which alleviate the responsibility and willing to pay back of borrowers. The explanation of type D listings is the same. Second, when rate and duration is large, the possibility of default is large because borrowers are not willing to pay back. Third, males are more likely to default. The result is the same as previous research [19]. Forth, wage owners are more likely to default because they have more living pressure than other occupations and the average financial situation is not good.

2) Rationality of herding when choosing listings

The herding index is not present in Table 5 because the p value is very high. To validate such result, we incorporate independent variable: the number of bidders, which is the same method of precious research, and the p value is still high. Thus we believe when lenders choose listings, the rationality of herding behavior is not certain.

Though herding in the listing is irrelevant to default, herding in the borrowers is significantly relevant to default. Thus considering the two aspects of herding behavior both, we conclude that herding behavior is irrational. The result is the same as Chen D *et al.* (2014)

VI. Conclusion

We study herding behavior in P2P lending in this paper. We try to solve two problems: whether there exists a herding behavior and whether it is rational. We divide herding behavior into two aspects: herding when lenders choose borrowers and herding when lenders choose listings. We assume there exists herding behavior and such herding behavior is irrational. Data is from PPDai, the largest P2P lending platform in China. Empirical study proves that there exists herding behavior in P2P lending and when we examine the rationality of herding, we define herding index to express the extent of herding. Consistent with previous research, we find herding in P2P lending in China is irrational. We believe the paper contribute study of herding in some aspects: first we deep discuss the herding and divide it into two aspects, second we define herding index based on Gini coefficient to express the extent of herding, third we use new dataset from China and enrich the empirically study of herding.

We believe that some further work can possibly enrich the research. First, laboratory experiment can be introduced to cross-validate the result because laboratory experiment suits the case when analyzing behavioral intention. Second, more data from different P2P lending websites should be tested.

Acknowledgment

This paper is supported by professor Tu giving a lot of specific instructions and help. This paper is sponsored by Humanities and Social Sciences of Ministry of Education Planning Fund “Research on the Decision-making Behavior of Investors and Borrowers and Risk Control Strategy in P2P online lending” (Project No. 15YJAZH066), Beijing Social Science Fund Project “Research on the Innovation Mode of Zhongguancun Small and Medium-sized Enterprises - Based on the View of Crowdsourcing Mechanism in Network Community” (Project No. 13JGA004) and scientific research innovation fund for graduate students of Central University of Finance and Economics (Project No. 201526). Thank all of them very much.

References

[1] E. Lee and B. Lee, "Herding behavior in online P2P lending: An empirical investigation," *Electronic Commerce Research and Applications*, 2012, vol. 11, no. 5, pp. 495-503.

[2] J. Light, "Is' Peer-to-peer'lending worth the risk," *Wall Street Journal*, 2012.

[3] M. Lin and S. Viswanathan, "Home bias in online investments: An empirical study of an online crowdfunding market," *Management Science*, 2015.

[4] J R. Magee, "Peer-to-peer lending in the United States: Surviving after Dodd-Frank," *NC Banking Inst*., 2011.

[5] L. Puro, J. E. Teich, H. Wallenius *et al*., "Borrower decision aid for people-to-people lending," *Decision Support Systems*, 2010, vol. 49, no. 1, pp. 52-60.

[6] S. Li, J. Qiu, Z. Lin *et al*., "Do borrowers make homogeneous decisions in online P2P lending market?" *An Empirical Study of PPDai in China*, 2011, pp. 1-6.

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

[8] J. E. Stiglitz and A. Weiss, "Credit rationing in markets with imperfect information," *The American Economic Review*, 1981, vol. 71, no. 3, pp. 393-410

[9] B G A. Akerlof, "The market for "lemons": Quality uncertainty and the market mechanism," *Quarterly Journal of Economics*, 2013, vol. 84, no. 3, pp. 488-500.

[10] S. Freedman and G Z. Jin, "The signaling value of online social networks: Lessons from peer-to-peer lending," *Nber Working Papers*, 2014.

[11] R. Iyer, A. I. Khwaja, E. F. P. Luttmer *et al*., "Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending?" *AFA 2011 Denver Meetings Paper*, 2009.

[12] 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*, 2013, vol. 59, no. 1, pp. 17-35.

[13] J. Duarte, S. Siegel, and L. Young, "Trust and credit: The role of appearance in peer-to-peer Lending," *Review of Financial Studies*, 2012, vol. 25, no. 8, pp. 2455-2484.

[14] J. Michels, "Do unverifiable disclosures matter? Evidence from peer-to-peer lending," *The Accounting Review*, 2012, vol. 87, no. 4, pp. 1385-1413.

[15] S. E. Asch, "Studies of independence and conformity: I. A minority of one against a unanimous majority," *Psychological Monographs*, 1956, vol. 70, no. 9, pp. 1-70.

[16] U. Simonsohn and A. Dan, "When rational sellers face nonrational buyers: Evidence from herding on eBay," *Management Science*, 2008, vol. 54, no. 9, pp. 1624-1637.

[17] Y. Komarova and L. Gonzalez, "Competition against common sense: insights on peer-to-peer lending as a tool to allay financial exclusion," *International Journal of Bank Marketing*, 2015.

[18] S. Bikhchandani, "Welch I. a theory of fads, fashion, custom and cultural change as informational cascades," *Ivo Welch*, 1992, vol. 100, no. 5, pp. 992-1026.

[19] D. Chen, F. Lai, and Z. Lin, "A trust model for online peer-to-peer lending: a lender’s perspective," *Information Technology and Management*, 2014, vol. 15, no. 4, pp. 239-254.

[20] D. Liu, D. Brass, Y. Lu *et al*., "Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding," *Mis Quarterly*, 2015, vol. 39, no. 3, pp. 729-742.

[21] B. Luo and Z. Lin, "A decision tree model for herd behavior and empirical evidence from the online P2P lending market," *Information Systems and e-Business Management*, 2013, vol. 11, no. 1, pp. 141-160.

[22] R. J. Kauffman, X. Li, Y. Au *et al.*, *Payoff Externalities, Informational Cascades and Managerial Incentives: A Theoretical Framework for IT Adoption Herding*, 2003.

Xiangyu Wang was born in Shandong province in 1993. He received the bachelor degree in computer science and technology from Central University of Finance and Economics. He is a master degree candidate at Central University of Finance and Economics. His research interests include internet finance and social network.

Yan Tu is a professor of Central University of Finance and Economics. Her research interests include: the electronic commerce, social network, knowledge management and internet finance.