

Who Can Get Money?

Evidence from the Chinese Peer-to-Peer Lending Platform

Qizhi TAO^a, Yizhe DONG^{b1}, Ziming LIN^a,

a: School of Finance, Southwestern University of Finance and Economics, Chengdu, 611130, China

b: Business School, University of Aberdeen, Dunbar Street, Aberdeen, AB24 3QY, UK

Information Systems Frontiers (forthcoming)

Abstract

This paper explores how borrowers' financial and personal information, loan characteristics and lending models affect peer-to-peer (P2P) loan funding outcomes. Using a large sample of listings from one of the largest Chinese online P2P lending platforms, we find that those borrowers earning a higher income or who own a car are more likely to receive a loan, pay lower interest rates, and are less likely to default. The credit grade assigned by the lending platform may not represent the creditworthiness of potential borrowers. We also find that the unique offline process in the Chinese P2P online lending platform exerts significant influence on the lending decision. We discuss the implications of our results for the design of big data-based lending markets.

Keywords: peer-to-peer (P2P) lending; Fintech, offline authentication; listing outcomes; information asymmetry; China

* Corresponding author: Yizhe DONG, Business School, University of Aberdeen, Dunbar Street, Aberdeen, AB24 3QY, UK, Tel.: +44 1224272167; Email: yizhe.dong@abdn.ac.uk.

Who Can Get Money?

Evidence from the Chinese Peer-to-Peer Lending Platform

1. Introduction

Online peer-to-peer (P2P) lending has recently emerged as a new form of loan origination for the credit market. According to Lin *et al.* (2009) and Bachmann *et al.* (2011), it is defined as peer-to-peer unsecured lending between lenders and borrowers through online platforms without the involvement of financial institutions. This type of lending marketplace is designed to supplement traditional bank lending in order to meet the small-loan needs of individuals and small-to-medium enterprises (SMEs), which often encounter difficulty in borrowing money from traditional lending institutions.

Since the first online P2P lending platform, Zopa, was established in the UK in 2005, numerous P2P lending platforms have emerged all over the world, such as the Prosper and Lending Club in the US, isePankur and Auxmoney in Europe, SocietyOne in Australia, and Renrendai and CreditEase in China. Concurrently this market has grown exponentially². Within these platforms borrowers generally make their personal and credit information – including a description of the reason for borrowing, current financial situation and demographic characteristics – available to potential lenders as the main indicators of creditworthiness. The platform can quickly assess and assign risk grades through the use of information technology. Instead of relying on delegated monitoring with banks as intermediaries, lenders have the opportunity to process the financial and personal information

² In the UK, the peer-to-peer lending industry lent more than 2.2 billion GBP in 2015 and this number is more than twice that of the previous year, according to the figures released by the Peer-to-Peer Finance Association (P2PFA). In the US, loans originated by P2P lenders reached 20 billion USD by 2015, again double the amount of the previous year. In the Chinese P2P market in 2015, lending volume was about 982.304 billion CNY, a rise of 388% on the previous year's figures. The number of P2P platforms now stands at 2595 as reported by WDJZ.COM.

provided by the borrowers and directly offer a loan that meets their investment criteria. P2P lending companies have taken a fee for providing the match-making platform and credit-checking the borrower.

An important function of lending markets is to screen borrowers and allocate credit efficiently based on the borrower's creditworthiness. However, the market can break down due to the problem of information asymmetry that the borrowers and lenders may face (Stiglitz and Weiss, 1981). Traditionally, as financial intermediaries, banks play a significant role in reducing information asymmetries because they are thought to have the financial expertise and extensive experience to evaluate the creditworthiness of borrowers and allocate capital accordingly (Diamond, 1984). In the P2P online lending market, transaction costs are reduced by eliminating expensive intermediaries, but as non-financial experts dominate in this pseudonymous online lending market, the problems associated with information asymmetry, particularly adverse selection, may become more severe than those that emerge in the traditional lending market (Lee and Lee, 2012). How to mitigate information asymmetry between borrowers and lenders is critical for the online lending market's long-term success. Some research suggests that information asymmetry can, to some extent, be alleviated by the disclosure and analysis of the borrower's soft information (e.g. Herzenstein and Andrews, 2008; Iyer *et al.*, 2009; Dorfleitner *et al.*, 2016) and friendship and group networks (e.g. Chen *et al.*, 2009; Lin, 2013; Liu *et al.*, 2015; Bellefamme *et al.*, 2015).

The problems of information asymmetries are even more pronounced in the Chinese P2P online lending market. Unlike the case in the US and other western countries in which the credit scores of borrowers are directly provided by specialised and independent credit rating agencies, Chinese online platforms use their own credit rating system to assess the creditworthiness of borrowers, based on the information submitted by the borrowers

themselves³. Because of the lack of any official and trustworthy credit records of borrowers, false or inaccurate information can easily pass through the online verification process. Therefore, the lenders may not fully trust the online credit assessment and verifications provided by Chinese P2P platforms.

In order to enhance the trust between borrowers and lenders – and improve market efficiency –offline verification mechanisms have been developed and adopted by some Chinese online P2P lending platforms, such as Renrendai.com. They use offline investigation tools such as physical site visits and third-party referrals to verify the borrowers' information and assess their credibility. The combination of online and offline authentication processes in the Chinese P2P lending market provides an excellent and unique opportunity for studying the role played by offline processes in electronic marketplaces. Therefore, we use a large sample of listings from one of the largest Chinese online P2P lending platforms, Renrendai.com, to examine how the offline process impacts on lenders' financing decisions. Along with a focus on the role of offline authentication processes in the P2P lending market, we also investigate how borrowers' hard financial information, soft information and loan characteristics impact on the loan outcomes, measured by the probability of successful funding, interest rates of loans, and the probability of loan default.

As mentioned above, we find that borrowers who own cars or those earning higher incomes are more likely to receive a loan, pay lower interest rates, and are less likely to default. Our empirical results indicate that some lenders do not perceive that the credit grade provided by the lending platform is a reliable indicator of the borrower's creditworthiness. We also find an inversely *U*-shaped relationship between interest rates and the probability of funding success. Moreover, we find that gender, educational and marital discrimination exists

³ The People's Bank of China (the central bank) has started the collection of credit history of individuals and entities and calculating of credit scores, but the credit information is not shared with non-banking institutions, such as P2P lending platforms.

in the market. In addition, our results indicate that a screening process by means of the third party's commitment to screen borrowers improves their (the borrower's) access to credit. Also, the additional transaction costs raised by the offline process can be compensated by a lower interest rate charged on their borrowing. Finally, we find that the offline authentication process helps to reduce the taste-based discrimination in the P2P lending market, and that borrowers' financial information has no impact on the probability of funding success for offline authentication listings.

Firstly, this study contributes to the growing literature on alternative finance by exploring the rapidly developing P2P lending sector from the largest emerging market, China. Most existing studies focus on the leading US P2P lending platform, Prosper, which has made its data publicly available (e.g. Freedman and Jin, 2008; Weiss et al. 2010; Pope and Sydnor, 2011; Ravina, 2008; Lin et al. 2013; Iyer et al. 2016) and crowdfunding, another type of alternative financing activities (e.g. Belleflamme et al. 2015; Colombo et al. 2015; Cumming and Johan, 2016; Cumming and Zhang, 2016). Secondly, we extend asymmetry information theories by examining and explaining the effect of offline authentication processes on the P2P online lending market. The findings on the role of offline processes provide insights into a possible future trend in the online lending market. Thirdly, our paper also contributes the literature on big data analytics. Big data analytics helps organisations and individuals in more informed decision-making and now drives almost every aspect of our modern society such as business intelligence (Chen et al. 2012), knowledge management (Hota et al. 2015), public sector (Klievink, et al. 2016), and healthcare (Kankanhalli et al. 2016). With the help of large volumes of data from a variety of sources, this research explores the main determinants of funding success, loan interest rate and default rate in order to provide practical guidance to borrowers in P2P online lending communities. Our findings shed light on which listing strategy is more effective and efficient for borrowers in successfully funding their request and

which lending strategy is more effective for lenders in reducing default risk. Fourthly, we provide comprehensive background information on the history, recent development and lending mechanisms of the Chinese P2P lending platforms and afford interested parties a better understanding of this new and fast-growing P2P lending market. Last but not least, our study is related to crowdsourcing literature. Some prior studies focus on the roles and methods of crowdsourcing for disaster management (Poblet et al., 2017), the factors influencing the crowdsourcing decision (Thuan, et al. 2016) and designing a novel crowdsourcing intermediary system (Yuan and Hesieh, 2016). However, we provide some insight into the crowdsourcing process and outcome by conducting an analysis on a peer-to-peer online lending platform, one of the important forms of crowdsourcing.

The remainder of the paper is organised as follows. Section 2 reviews the related empirical literature on P2P lending. Section 3 presents the development of the Chinese P2P lending market and the lending process of the Renrendai platform. While Section 4 describes the data made available by Renrendai.com, we develop and analyse the empirical model and discuss the empirical results in Section 5. Finally, Section 6 concludes the paper.

2. Related Literature Review

In response to the growing popularity of P2P lending, a large amount of research in this area has emerged in recent years. Some studies investigate the reasons for – and economic impact of – the emergence of online P2P lending. Hulem and Wright (2006) study the first online lending platform – Zopa – and suggest that the emergence of online P2P lending is a direct response to social trends and a demand for new forms of financial services in the new financial technology age. Wang *et al.* (2009) find that P2P lending marketplaces could increase individual interactions and provide efficient information transfer, and thus perhaps create a more efficient market. Culkin *et al.* (2016) suggest that online peer-to-peer lending

(and crowdfunding) has some advantages over more traditional sources of entrepreneurial finance and can facilitate start-up and small-firm growth. Fraser et al. (2015) also point that P2P lending may contribute to filling gaps in the supply of banking finance and may facilitate the flow of information among the borrowers and lenders, however, the exact role of the fund providers in the context of monitoring and governance remains somewhat unclear. Mild et al. (2015) and Guo et al. (2016) proposed some models to assess the credit risk of the online borrowers. Most of the previous studies have mainly focused on identifying the factors that determine the funding success, the interest, and/or the default rate. On a meta level, the factors can be classified into two broad categories: (1) hard information, e.g. borrowers' credit scores and financial situation, and (2) soft information, e.g. demographic characteristics, social networks, photographs and descriptive text.

Klaft (2008) analyses the data retrieved from Prosper and shows that borrowers' credit ratings and verified bank account information significantly impact on the probability of successful funding. They also find that credit rating and debt-to-income ratios are key determinants for loan interest rates, but that verified bank account information and home ownership have almost no impact on interest rates. Similarly, Herzenstein and Andrews (2008) find borrowers' credit ratings significantly impact upon the likelihood of funding success. Iyer *et al.* (2009) examine the impact of borrowers' credit ratings on the interest rate of a funded loan; they find that lenders rely more on hard financial information to differentiate between borrowers within the higher credit rating categories, while delinquencies and other forms of soft data are more informative within the lower credit categories.

Some studies also exploit data on the personal characteristics of borrowers to test theories of statistical and taste-based discrimination in the P2P lending market. Duarte *et al.* (2012), Pope and Sydnor (2011) and Ravina (2008) focus on facial attributes such as race and beauty of the borrowers, addressing the literature on racial bias and the beauty premium

based on the data from Prosper. Duarte *et al.* (2012) and Ravina (2008) both find that borrowers with more positive characteristics – like trustworthiness and beauty – have a higher probability of having their loans funded and also tend to default on their loans less often. Pope and Sydnor (2011) present evidence for taste-based discrimination against blacks, the elderly and the overweight; these borrowers are less likely to secure loans and more likely to pay higher interest rates if they do. Loureiro and Gonzalez (2015) find that more attractive and financially successful borrowers of the same gender as lenders are less likely to have their loan requests filled and suggest that interpersonal competition significantly influences online lending decisions.

There is much evidence that other soft information such as social networks and listing descriptions play significant roles in online P2P lending and help in mitigating asymmetric information. Berger and Gleisner (2009) analyse all listings on Prosper.com between November 2005 and September 2007 and find that the designated group leaders can act as intermediaries; this helps to reduce information asymmetries and can lead to the group membership achieving significantly reduced borrowing costs. Similarly, Chen *et al.* (2009) reveal that the borrowers' general group social capital (i.e. group membership) improves funding success and reduces interest rates and default probability, while Freedman and Jin (2008) find that loans endorsed by friends, and bids, show fewer missed bids have fewer missed) payments and higher rates of return. Using a sample of Prosper listings, Lin *et al.* (2013) examine the role that friendship networks played in the outcomes of loans on the P2P market. They find that the friendship can serve as an informational boost to a borrower's credit quality, increase the probability of successful funding, and reduce the interest rate of loans and default risk on funded loans. Liu *et al.* (2015) also find that the borrowers' friends, particularly close offline friends, are more likely to bid on the listings.

In addition, Dorfleitner *et al.* (2016) study the relation between the soft information

derived from the listing description and the lending outcomes based on data from two leading European P2P platforms. They find that spelling errors, text length and the presence of social and emotional keywords in the description text significantly negatively influence the likelihood of funding success, but are not significantly related to the default probability. Iyer *et al.* (2016) show that lenders rely on soft information such as the characteristics of listing text to infer borrowers' creditworthiness, particularly when evaluating lower-quality borrowers.

3. Institutional Background

3.1 Chinese P2P lending market

Since the first Chinese online P2P lending platform, ppdai.com, started operating in 2007, this lending market has grown at a phenomenal rate and is currently the largest in the world. According to a report in China News⁴, as of the end of 2015, the sector was home to more than 2600 platforms and topped about 982.3 billion yuan or \$149 billion, soaring from 253 billion yuan in 2014 (see Figure 1 for more details). This rapid growth has been driven by the huge demand from individuals, households, small and micro-enterprises for access to Chinese finance; and by the supply of funds from retail investors. P2P loans typically carry hefty interest rates in the 10% to 20% range – that is three to six times higher than the bank deposit rate. According to data disclosed by WDZJ.com, average interest rates offered by the Chinese P2P platforms have significantly declined to around 13.29% in 2015 from nearly 20% two years ago.

[Insert Fig. 1 here]

In China, the current regulatory requirements for P2P lending platforms are relatively

⁴ <http://www.ecns.cn/business/2016/01-02/194408.shtml>

light. There has been no formal regulation from the central bank (People's Bank of China) or banking regulator (Chinese Banking Regulatory Commission) until recently. The Chinese authorities introduced their first major guidance policy on internet finance in July 2015⁵. According to the guidance, P2P lending providers should consider themselves as *information* intermediaries rather than *credit* intermediaries, and need to establish a third-party depository account for customer funds with a qualified banking institution (Deer, *et al.*, 2015). This loose regulatory environment has proven to be a double-edged sword; it allowed the rapid growth of the P2P lending market, providing new entrepreneurial opportunities and bringing convenience to the borrowers, but loose regulation also made it easy for bad debtors to defraud unwary investors. For example, in early 2016, nearly a million investors lost \$7.6 billion through Chinese online 'Ponzi-like schemes' orchestrated by Ezubao, the country's largest peer-to-peer lending service⁶. In response to the discovery of Ponzi schemes in several Chinese P2P platforms, tighter regulations on P2P lending are expected to promulgate soon.

3.2 Renrendai.com

Our analysis focuses on Renrendai.com which is a leading peer-to peer online lending platform in China. Renrendai was founded in 2010, and, by the end of 2015, it had confirmed more than 220,000 loans with a total lending amount of over 13.21 billion RMB (Chinese Currency) (see Figure 2 for more details).

[Insert Fig. 2 here]

In order to be considered for a loan from the Renrendai online lending platform, a potential borrower must file an application by providing his or her national ID number and other personal information so that Renrendai can authenticate the applicant's identity, mobile

⁵ People's Bank of China and 10 other Ministries and Commissions jointly issued "Guiding Opinions on Promoting the Healthy Development of Internet Finance" in 2015.

⁶ See ft.com. Link: <http://www.ft.com/cms/s/0/4ca011f4-c88f-11e5-a8ef-ea66e967dd44.html#axzz4Di9nuZIy>

phone, address, employment status, and income. In the listing, the borrowers must describe the purpose of the loan, the requested loan amount, the interest rate they are willing to pay, the duration of the loan request, and any other information (including a photo and description of themselves) that they feel may help attract lenders to fund the loans. In addition, the platform also encourages potential borrowers to provide evidence about their home and car ownership status, education level, professional qualification, marital status, residential address, and other evidence/certifications that may help improve their credit scores and financing limits. Unlike many online lending platforms in Western countries which can pull the borrower's credit history and score from professional credit rating agencies, Renrendai relies on its own credit rating model to assess borrowers' credibility. Renrendai assesses and verifies the borrowers' submitted information, also checking their historical credit information if they have the previous history of borrowing from Renrendai. A credit score is then assigned to each borrower. Based on the credit score, the lending platform further classifies every loan application into seven risk-grades starting with the safe *AA*, followed by Grades *A* through *E*, and finally *HR* (high risk).

Because of the relative lack of trustworthy credit records and sophisticated credit rating models, lenders do not always fully trust the credit scores and grades posted by Renrendai. Therefore, in addition to the aforementioned pure online credit assessment system, the platform has also developed an offline credit assessment process to help lenders and borrowers establish trust and increase the number of successful applications. Renrendai seeks to establish a direct offline relationship with borrowers by first directing them to an offline branch managed by its sister company, Ucredit (www.ucredit.com), to submit application materials for the loan as part of its "offline authentication" process. When Ucredit receives loan applications, it screens applicants by undertaking physical site visits to verify borrowers' information and then forwards the qualifying borrowers to Renrendai. Any offline referrals

from Ucredit are automatically classified as Credit Grade A when they are listed in the online platform. This offline authentication process has grown in popularity due to its potential to alleviate asymmetric information between borrowers and lenders. When Renrendai launched in 2010, it adopted a pure online operating model similar to that of the Lending Club. The platform now tends to rely much more on the offline operating model after integrating with offline lending service Ucredit in 2012. Approximately 89% of the loan amounts sourced in 2015 are recorded as having completed offline authentication (See Figure 3 for more details on loan mix). Moreover, some potential borrowers also have an approved third-party credit guarantee company backing their bid and seeking raise funds on their behalf via Renrendai⁷.

[Insert Figure 3 here]

As with borrowers, potential lenders must provide their national ID number, mobile phone number and bank information for Renrendai to verify their identity. Lenders can browse listings pages which include all of the borrowers' information described above, and the percentage funded. After assessing this information, a potential lender can bid on a listing by specifying the amount of money he or she would like to contribute. The minimum bid requirement is 50 RMB. Most lenders limit their bids to a small amount on each loan and diversify their funds across many different loans. A listing will close once the total amount bid exceeds the borrower's request.

Renrendai charges fees for providing the match-making services between borrowers and lenders. There is no fee for posting a listing, although borrowers need to pay a closing fee

⁷ The smart-money is an intelligent lending feature that allows investors to automatically allocate their money across the listings by selecting their desired risk appetite and saving plan objectives. However, Renrendai stopped the smart-money lending mode at the end of 2012.

when their loans are originated. The fee ranges from 0% to 5% depending on credit grade. In addition, borrowers are also charged a 0.3% annual serving fee based on the current outstanding loan principal. Except for a small cash transaction fee, the platform does not generally charge servicing fees for lenders. All Renrendai loans are uncollateralised, up to 36 months, and fully amortised with fixed interest. Loan size can range from 3,000 RMB to 500,000 RMB. Renrendai has also set up a risk reserve fund which is designed to cover borrowers' missed and late loan repayments; this is constantly topped up using part of the fee paid by each borrower⁸. In the event of any missed payments by borrowers, the reserve fund makes the expected payments to lenders instead. If a repayment is overdue, the platform makes several attempts to recover the loan, including emailing, text messaging, and calling the borrower. After payment is overdue by two months or more, a collection agency takes charge of collecting the default. If money is subsequently recovered from borrowers, this is paid back into the reserve fund.

4. Data and variables

Our sample comprises all listings that sought funding on Renrendai.com between 1 January 2013 and 31 December 2015. After dropping listings with missing variables, we arrive at a final sample consisting of 499,180 listings; of these, 197,743 listings had been successfully funded. Each listing and loan in our data set is associated with a number of variables that include all of the information seen by the lender when they made their lending decisions, as well as the outcome and performance of the listings/loans. These variables fall into six categories: (1) dependent variables to measure the outcome of a listing and the performance of a loan; (2) credit profile and financial information; (3) information describing

⁸ The Shanghai Branch of China Merchants' Bank has been providing a custodian service for the risk reserve fund. of Renrendai.com and issues monthly report on its fund flows.

the specific features of a listing/loan; (4) demographic information; (5) listing types, and (6) other control variables. A complete list of all variables obtained from Renrendai is summarised in Appendix A and their descriptive statistics are reported in Tables 1 and 2.

[Insert Table 1 here]

[Insert Table 2 here]

We construct three dependent variables in this empirical study to measure the outcomes of listings or loans. The first dependent variable (*LISTSUC*) is a binary variable to indicate whether the listing attracted enough bids and generated a loan. The second dependent variable is the interest rate charged on a funded loan (*LRATE*). The last dependent variable is a dummy variable to indicate whether a successful listing defaults (*LDEF*). According to Table 2, 40.4 % listings have been successfully funded. The average interest rate of the loans is 12%; this rate has varied over time, ranging from 8% to 24%. The default rate of loans is small and only 2,214 loans have been labelled as defaults in our sample.

Many previous studies show that the hard credit and financial information of borrowers can impact upon their creditworthiness and affect lenders' online lending decisions (Herzenstein and Andrews, 2008, Berger and Gleisner, 2009; Iyer *et al.*, 2016). The credit rating grade is one of the most common indicators used by a lender to assess a borrower's creditworthiness when making his or her lending decision. The Renrendai awards seven grades for borrowers, which range from *AA* to *HR*. To measure the impact of this factor, we include a full set of dummy variables for each credit grade (*CRED*). Higher credit grades indicate a higher probability of the borrower paying back the loan, so we would expect the variable to have a positive influence on funding success. Meanwhile, since a higher credit score usually results in investors reducing their risk expectations, it should have the effect of decreasing the risk premium requested by investors; therefore the variable should have a negative impact on the loan interest rate.

We also factor other hard credit information on the listing such as borrower's income levels (*INC*) and credit line (*CREDITL*) into our regressions to proxy for the borrower's financial standing. In addition, following Duarte *et al.* (2012), we include variables to identify whether the borrower is a home (*HOME*) or car (*CAR*) owner, factors that can be considered indications of wealth or creditworthiness. According to Table 2, Renrendai is characterised by a dominance of applications in the lowest rating category *HR* (60.6%) and a scarcity of applications with the 'medium' Grades *B-E* (2.3%). Of the borrowers, 37% and 0.1% have been classified as good, rating *A* and *AA*, respectively. As expected, funded listings tend to have borrowers with better credit grades; 91.2% of the funded listings are Credit Grades *AA* and *A* while only 5.7% loans originate from successfully funded *HR*-listings. Regarding the borrowers, 1.3% of their monthly incomes are below 2,000 CNY; 84.3% claim that their monthly incomes range between 2,001 and 20,000; and 14.3% of borrowers' monthly incomes are over 50,000.

The characteristics of the loan application – for example, the interest rate that the borrower is willing to pay, the loan amount requested, and the loan length – are directly related to the return of the investment and are essential elements to be considered when the lender makes his lending decision. When a potential borrower creates a listing, he sets the interest rate he is willing to pay. The interest rate offered will to some extent reflect the borrower's self-assessment of the fairness of his individual risk premium. A higher interest rate offered by the borrower can increase the listing's attractiveness for more lenders and stimulate them to bid on the loan request; however, as the interest rate increases, the default risk increases as well which may reduce lenders' profitability and subsequent incentive to bid on the listing. Therefore, we expect an inversely *U*-shaped relationship between the interest rates and the probability of funding success. Following Lin *et al.* (2013), we include the interest rate (*LISTRATE*) and its quadratic term (*LISTRATE2*) in our regressions to capture

this relationship. We also incorporate two additional variables to measure the loan amount requested (*LSIZE*) and loan length (*LLENGHT*) which are related to risks and affect the lenders' lending decisions (Herzenstein and Andrews, 2008; Chen *et al.*, 2009). Comparing listings and loans, the average amount requested for listings is 62,278 CNY while the average loan size is 61,871yuan. The average loan length is higher for the successful listings (27.9 months) than for all listings combined (21.27 months); it appears that lenders prefer listings with a larger loan size and a longer duration. The average interest rate that borrowers are willing to pay is 13%; this is slightly higher than the average contract rate of 12%.

Some studies have found that a lender's subjective assessment of a potential borrower's creditworthiness correlates highly with the borrower's demographic characteristics (e.g. Ladd, 1982; Carr and Megbolugbe, 1993; Herzenstein and Andrews, 2008). Therefore, we would expect that the demographic information provided by the borrower is important in determining whether the listing gets funded and how much interest is charged. Following Pope and Syndnor (2011) and Barasinska and Schäfer (2010), we include an extensive set of variables to capture the borrower's gender (*FEMALE*), age (*AGE*), educational background (*EDU*) and marital status (*MAR*). Of all listings, only 18.5% were female; the average age of borrowers was 33 years; 32.9% of borrowers had the lowest level of education (high school level education or lower), 44.3% had a college degree, and 21.5% had attained a Bachelor's degree, while only 1.4% of borrowers claimed to have an advanced degree (master's degree or above). Of the potential borrowers, 56.9% were married, 36.2% were unmarried and 0.7% were divorced or widowed.

One of the major problems of P2P lending is asymmetric information (Freedman and Jin, 2008; Lin *et al.*, 2013). This problem is more pronounced in the Chinese P2P lending market due to the lack of reliable and comprehensive credit information. Chinese P2P lending platforms have developed a more diverse set of lending models compared to the case in the

US and other countries in its response to the challenges of limited reliable credit information, the nature of investor risk, and different borrower segments (Deer *et al.*, 2015). Renrendai has adopted three lending models – pure online, offline authentication and third-party referral lending models. The major difference between these lending models is the way they screen potential borrowers' information. The first lending model relies purely on an online screening process, and not all borrower-provided information is verified. The other two models verify borrowers' information via physical site visits in the first instance and third-party credit guarantee companies in the second. These additional screening efforts could significantly increase the reliability of information provided by the borrowers. Moreover, borrowers with better credit quality are more inclined to use offline processes to signal their creditworthiness to outside lenders. Therefore, these offline lending models can mitigate the problem of adverse selection, and we expect that offline authentication (*OFFLINE*) and third-party referral (*GUAR*) listings would more likely to attract funding at lower interest rates than would pure online listings (*PURE*). Among all the listings in the sample, there are 308,383 (63%) pure online listings, 158,620 (32.5%) offline authentication listings and 22,177 (4.5%) third-party referral listings. Almost all offline authentication listings (158,438/158,620 = 99.88%) and third-party referral listings (98.7%) are able to attract sufficient bids and secure funding; however, only 5.65% pure online listings had been successfully funded.

We also include a number of control variables in our analysis. Some investors may favour credit bids on certain special loan categories, leading to a higher probability of a successful funding and a higher interest rate (Lin *et al.*, 2013; Dorfleitner *et al.*, 2016). We classify the intended purposes of loans into seven categories: debt consolidation, buying a house, bill payments, business loans, automobile loans, home improvement loans, medical loans, education loans, wedding loans and other loans. We include a full set of dummy variables for each loan category (*LC*) to control for investors' preference. We see that 48.1%

of all listings are debt consolidation, followed by bill payments (19.8%), home improvement (10.3%), and business loans (8%). We further include a set of dummy variables to control for the geographical region (*REG*) in which the loan applicant's residence is located⁹. In addition, we include dummy variables to indicate the borrower's employment status (*EMP*) and employment length (*EMPYEAR*); these factors can indirectly convey information about his or her financial situation and debt-servicing capabilities while also influencing both funding success and interest rate (see Iyer *et al.*, 2016). Finally, following Dorfleitner *et al.* (2016), we include an additional control variable *LTEXT*, which comprises the number of words included to measure the length of the descriptive text.

5. Empirical analysis and results

5.1 Empirical models

This study applies regression analysis in order to determine the factors that impact the probability of a listing being funded, as well as the interest rate and default probability of funded loans as the outcomes of the lending transaction in the P2P lending platform. We estimate the following model three specifications:

$$\begin{aligned} \text{Probability}(FSUC = 1|x) &= \alpha_0 + \alpha \text{FinInfo} + \beta \text{LoanChart} + \chi \text{PersonalChart} \\ &+ \theta \text{ListType} + \delta \text{OtherCon} + \varepsilon_i \end{aligned} \quad (\text{Eq.1})$$

$$\begin{aligned} E(LRATE|x, FSUC = 1) &= \phi_0 + \phi \text{FinInfo} + \varphi \text{LoanChart} + \gamma \text{PersonalChart} \\ &+ \eta \text{ListType} + \kappa \text{OtherCon} + \varepsilon_i \end{aligned} \quad (\text{Eq.2})$$

and

⁹ We divide China into six economic regions plus an extra category made up of four municipalities. The six regions are the centre and south (Henan, Hubei, Hunan, An'hui and Jiangxi), the east coast (Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan), the north (Inner Mongolia, Hebei and Shanxi), the northeast (Jilin, Liaoning and Heilongjiang), the southwest (Sichuan, Guizhou, Yunnan, Guangxi and Tibet), and the northwest (Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang). The four municipalities are Beijing, Shanghai, Tianjin, and Chongqing.

$$\begin{aligned} \text{Probability}(LDEF|x, FSUC = 1 \& ONLINE = 1) = & \mu_0 + \mu_{FinInfo} + \theta_{LoanChart} \\ & + \rho_{PersonalChart} + \zeta_{OtherCon} + \varepsilon_i \end{aligned} \quad (\text{Eq.3})$$

where *FSUC*, *LRATE* and *LDEF* are the dependent variables which measure whether the listing had been successfully funded, the interest rate charged for the loan, and whether the loan defaulted respectively. *FSUC* and *LDEF* are both binary variables; hence, we use *Probit* regressions to estimate the probability that a listing is funded successfully (Eq. 1) and also the probability that a funded loan defaults (Eq. 3)^{10,11}. Because the interest rate is bounded by zero and one, we use Tobit regression to estimate the equations 2. *FinInfo* denotes hard credit variables including credit grade dummies, credit line, income status and home and car ownership. *LoanChart* describes the specific features of a listing including the amount requested, loan size, and the interest rate the borrower is willing to pay. *PersonalChart* refers to personal characteristic variables including gender, age, educational background and marital status. *ListType* captures whether a listing belongs to pure online, offline authentication or third-party referral listings. *OtherCon* denotes a set of control variables including loan category, length of loan description, employment status, and borrowers' residential areas. Finally, *x* represents the information set about the listing, and ε_i is the random term. In the following sections, we assess how the above information impacts the lending outcomes in detail, based on our regression results.

5.2 Funding probability

Table 3 shows the *Probit* regression results for the model specifications with *FSUC* as a dependent variable (Eq.1). Models 1-4 incorporate exploratory variables *FinInfo*, *LoanChart*, *PersonalChart* and *LendType* separately, each together with the control variables. Model 5

¹⁰ In the sample, we find there are no defaults on the loans which were originated by the offline authentication and third-party referral lending models. All default loans come from the loans which were funded based on the pure online lending process. Therefore, we examine the determinants of default probability based on the sample which only includes the loans made by the pure online lending process.

¹¹ We also use Logit models to estimate the probability that a listing is funded successfully and a loan default is used as a robustness check. The Logit regression results are consistent with those from the *Probit* models.

represents the main funding probability model including all variables simultaneously.

As expected, the hard financial variables have a significant impact on funding probability. The results in Table 3 show that the coefficients on credit rating variables are all positive and highly significant, suggesting that the *HR* rated listings are less likely to be funded than other, higher credit grade listings. However, in contrast to the findings of Lin *et al.* (2013) and Ravina (2008), it is interesting to note that the funding probability does not monotonically increase with the credit rating. Some low credit grade listings (e.g. *E*) seem to have higher probability of being funded than some better credit grade listings (e.g. *C*). This is probably due to the fact that, in the absence of a reliable and official credit information environment, Chinese P2P lenders may not fully trust the credit grades provided and also rely on other information to make their lending decisions. In addition, we find that the borrower's monthly income has a significantly positive effect on the probability of funding success, suggesting that investors make their lending decisions on the basis of the borrower's income level. As the monthly income increases, all things being equal, the probability also increases. Car ownership has a significant positive effect on the borrower's likelihood of funding success, although the effect is less significant than that of good credit grades and higher income levels.

As discussed above, there may be a non-linear relation between the interest rate a borrower is willing to pay and the probability of funding success. In regressions 1 and 5, the coefficients on the level terms (*LISTRATE*) are statistically significantly positive, while the coefficients on the quadratic terms (*LISTRATE2*) become significantly negative. This constitutes an inversely *U*-shaped pattern, suggesting that a higher interest rate could increase the funding probability but, as the interest rate increases above a certain level, this effect may be swamped by the greater likelihood of default. High interest rates may indicate risky borrowers, who are less likely to be funded. With regard to the loan amount requested, we

find that the coefficients of *LSIZE* are negative and highly significant, suggesting that lenders tend to bid on small loans to diversify their risk. As the loan amount increases, the lenders' assessments of borrowers' ability to repay loans may be less favourable, reducing the probability of funding success (Herzenstein and Andrews, 2008). In addition, the results also show that the loan length has a significant and positive impact on the likelihood of getting a loan, and suggest that Chinese P2P lenders may prefer to bid on longer term loans in order to secure a high level of return for a longer period.

Apart from the hard financial information and loan characteristics, we find some evidence of discrimination in the Chinese P2P lending market. The results show that the female and older borrowers were more likely to be successful in their loan requests. The coefficients on education variables are all positive and highly significant. This result indicates that borrowers with 'better' educational backgrounds are more likely to have their loan funded. Similarly, the coefficients on *MAR2* are also positive and statistically significant, which suggests that lenders were more likely to deny loan requests from unmarried applicants compared to married applicants. It is worth noting that, although demographic characteristics – for example, gender, age, educational background and marital status – affect the likelihood of funding success, these effects are relatively small compared to the borrower's financial information and loan characteristics.

As we discussed above, some Chinese P2P lending platforms have developed a more diverse set of lending criteria in response to the challenges posed by the lack of reliable credit information on borrowers. Renrendai currently relies mainly on pure online (*PURE*), offline authentication (*OFFLINE*), and third-party referral (*GUAR*) processes to source borrowers and to verify borrowers' information. As shown in Table 3, the coefficients on *OFFLINE* and *GUAR* are highly significant and positive; this suggests that the method of verifying credit information plays a very strong role in the lender's decision-making process. The screening

efforts undertaken by the offline physical site visits and the third-party credit guarantee companies lead to reduced adverse selection problems and a significantly increased probability of a loan's successful funding. From the perspective of borrowers, high-quality borrowers may also be required to undergo an offline authentication process to signal their creditworthiness, which in turn alleviates asymmetric information and attracts lenders to bid on these listings.

[Insert Table 3 here]

5.3 Interest rates on funded listings

We next examine how borrowers' information influences the interest rate charged to the borrower. As with funding probability, we regress the funded loan's interest rate on the borrowers' related variables. Models 1-4 incorporate exploratory variables *FinInfo*, *LoanChart*, *PersonalChart*, and *ListType* separately, each combined with the control variables. Model 5 represents the main interest rate model including all variables simultaneously. The results of the *Tobit* regression are illustrated in Table 4. In general, the results are consistent with those for funding probability.

As shown in Table 4, the borrowers with the highest credit grades (*AA* and *A*) pay significantly lower interest rates than other lower credit grade borrowers, except the Grade *E*. Surprisingly, we find that the Credit Grade *E* borrowers seem to pay the lowest interest rates and the medium quality credit borrowers (*B* and *C*) pay the highest interest rates. The results indicate that Chinese P2P lenders may not perceive the credit grade provided by the platform as a good indicator of the creditworthiness of the borrower, and their investment decisions are also likely to be affected by other factors. As expected, higher income borrowers (*INC1-INC6*) pay significantly lower interest rates than the lowest income range (less than 1,000 RMB). However, the interest rates paid by the borrowers whose income levels are within the *INC1* to

INC6 ranges are not significantly different. In addition, we find that, on average, car-owning borrowers pay 10 basis points (or 0.1%) more interest rate than borrowers without cars, while home ownership has the opposite effect: having a home decreases the interest rate by about 0.1% compared to the base group.

The coefficients on *LSIZE* are negative and highly statistically significant, suggesting that the higher the loan amount requested, the higher the interest rate. As expected, we also find that the borrowing period has a significantly positive effect on the interest rate, suggesting that the longer the term to maturity, the higher the interest rate. This is probably due to the fact that, since longer period loans are generally subject to greater interest rates and credit and liquidity risks, additional risk premiums will be required.

Regarding borrower attributes, the coefficients on *FEMALE* and *AGE* are rather small and almost equal to zero in Model 5, suggesting that the borrowers' gender and age have little effect on the interest rates paid. However, the estimated coefficients on the education background are sensible and highly statistically significant. Compared to the base group of borrowers who have only completed secondary education, those who have college degrees pay 0.1% less interest, while those who have Bachelor's or higher degrees pay 0.2% less interest. Although we previously found that the single borrowers are more likely to have their loans funded, on average, they have to pay 0.1% more interest than groups of borrowers. Taken together, the effects of these personal characteristics are statistically significant but their economic magnitude is very small.

Finally, we find that the coefficients associated with *OFFLINE* and *GUAR* are negative and highly statistically significant. As shown in Model 5, the loans originated through the offline authentication process and the third-party referral have lower interest rates compared to the loans relying solely on the pure online process; by 1% and 1.8%, respectively. The results suggest that the P2P lending models are the most important factors affecting the

interest rate charged to the borrower, and that the use of the offline process (the offline authentication and third-party credit guarantee) can significantly reduce the asymmetric information.

[Insert Table 4 here]

5.4 Probability of default

The previous sections have shown how hard financial information, personal information, and loan features affect funding probability and the interest rate charged. It is also of interest to us to examine whether these characteristics have any impact on the probability of default. Our data cover all the loans originated between January 2013 and December 2015, under either repayments or default; these data show that all defaults come from the pure online-originated loans; therefore our analysis focuses exclusively on the pure online loans and excludes the offline-originated loans from the sample. Table 5 reports the *Probit* model estimations for the default probability.

For the credit rating variables, we find that the two extreme Credit Grades (*HR* and *AA*) are *less* likely to default while Grades *B* and *D* are *more* likely to default. The results suggest that credit grade is not a good predictor of default. Moreover, all coefficients on borrowers' income levels are statistically insignificant thus suggesting that this factor is not related to the default probability when all the other characteristics are taken into account.

As with funding probability, interest rate exhibits a non-linear relationship with default probability, suggesting that in the low-level interest rate range, increasing the interest rate reduces the default probability but, once the rate increases above a certain level, the high rates are associated with risky borrowers who tend to default on the loan. We also find that the short period loans are associated with lower odds of default; however, we do not find that loan size has a significant impact on the default probability.

The results also show that, all things being equal, female and older borrowers are more likely to default. The coefficients on the education dummy variables are positive and statistically significant, and indicate that borrowers with a higher education background may tend to take more risks and therefore have a higher default probability.

[Insert Table 5 here]

5.5 Pure online versus offline authentication processes

One of the most important findings in the previous sections is that the offline authentication processes could significantly alleviate the problem of asymmetric information and influence the lenders' lending decisions. It is therefore natural to expect that the impacts of borrowers' financial information, personal attributes and loan characteristics on the probability of funding success and the interest rate of funded loans are different depending on whether the authentication listings/loans are online or offline. To test these expectations, we re-estimate Equations 1 and 2 based on pure online and offline authentication sub-samples. The regression estimations are reported in Table 6. The second and third columns report the *Probit* regression results of the probability of funding success based on the sub-samples of offline authentication and pure online listings, respectively. The fourth and fifth columns report the *Tobit* regression results of the interest rate charged, based on the sub-samples of offline authentication and pure online listings, respectively.

As expected, our results show that the impacts of borrowers' financial information, personal attributes and loan characteristics on the probability of successful funding differ significantly between the pure online and the offline authentication listing groups. Estimates for the pure online sub-sample remain basically unchanged from the results based on the full sample; however, the estimates for the offline sub-sample have largely changed and many of them become statistically insignificant. We find that the coefficients associated with the

financial information variables are no longer statistically significant, suggesting that the borrower's income, car ownership and home ownership have no significant effect on the probability of funding success. Similarly, the coefficients associated with gender, education background and marital status become insignificant or less statistically significant. These results indicate that the offline authentication process could help to reduce taste-based discrimination which is often found in credit markets (see Ross and Yinger, 2004; Pope and Sydnor, 2011). We also find that an offline listing with a low interest rate is less likely to be funded but that, once the interest rate increases to a certain high level, the likelihood of funding success significantly increase. In addition, the results show positive relationships between the loan amount requested, loan length, and funding probability – but these effects are not statistically significant.

In contrast with the funding probability, we do not find a much significant difference in the impact of the borrowers' information on the funded loan's interest rate between the pure online and the offline authentication loan groups. The estimates for these sub-samples show nearly the same impact directions and statistically significant levels, and are highly consistent with our prior findings based on the whole sample. However, it is worth noting that the magnitude of coefficients associated with income-related variables in the offline authentication sub-sample is much smaller than that of the coefficients associated with the pure online sub-sample. The results suggest that the borrowers' income levels have a stronger impact on the interest rates of pure online loans than on those of offline authentication loans.

[Insert Table 6 here]

6 Conclusion

Online P2P lending is an unsecured lending activity between lenders and borrowers through online platforms without the involvement of financial institutions. It is designed to

supplement traditional bank lending in order to meet the small-loan needs of individuals and SMEs. In order to enhance the trust between borrowers and lenders –and improve market efficiency—an offline verification mechanism has been developed and adopted by some Chinese online P2P lending platforms. The combination of online and offline authentication processes in the Chinese P2P lending market provides an excellent and unique opportunity for studying the role played by offline processes in electronic marketplaces.

This paper uses a large sample of listings from one of the largest Chinese online P2P lending platforms, Renrendai.com, between 1 January 2013 and 31 December 2015, to study how borrowers' financial and personal information, loan characteristics and lending models affect the P2P loan funding outcomes measured by the probability of successful funding, interest rate of loans, and the probability of loan default.

This research sheds light on the role of the offline process on the online P2P lending market. Our results indicate that a screening process by means of the third party's commitment to screen borrowers improves the borrowers' access to credit, and the additional transaction costs raised by the offline process can be compensated by a lower interest rate charged on their borrowing. The offline process can reduce information asymmetries between borrowers and lenders. Moreover, we find that the offline authentication process helps to reduce the taste-based discrimination in the P2P lending market, and borrowers' financial information has no impact on the probability of funding success for offline authentication listings. One of the important implications emerging from this study is its suggestion that online P2P lending platforms adopt a more diverse lending models including both online and offline lending checking processes which will help to mitigate the problems of asymmetry information in P2P lending markets.

Our study makes contributions to the literature in the following four aspects. Firstly, this study contributes the growing body of literature on alternative finance by exploring the

rapidly developing P2P lending sector from the largest emerging market – China. Secondly, we extend asymmetry information theories by examining and explaining the effect of offline authentication processes on the P2P online lending market. The findings on the role of offline processes provide insights into a possible future trend in the online lending market domain. Thirdly, this research explores the main determinants of funding success and loan interest rate in order to provide practical guidance to borrowers in P2P lending communities. The findings shed light on which listing strategy is more effective for borrowers in successfully funding their request. Finally, we provide comprehensive background information on the history, recent development and lending mechanisms of the Chinese P2P lending platforms and afford interested parties a better understanding of this new and fast-growing P2P lending market.

Acknowledgments: We thank the Editors, Ram Ramesh and Raghav Rao, the guest editors, Douglas Cumming, Sofia Johan, and Denis Schweizer for their helpful and valuable suggestions. We are grateful for useful comments from the participants of the 2nd microfinance and rural finance conference in Aberystwyth. Qizhi Tao acknowledges support from the Fundamental Research Funds for the Central Universities (Grant No. JBK160921).

References

Accessed 10 May 2016

- Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Lehmann, M. & Tiburtius, P. (2011). Online peer-to-peer lending- a literature review. *Journal of Internet Banking and Commerce*, 16(2), pp. 1~18.
- Barasinska, N. & Schäfer, D. (2010). Are women more credit-constrained than men? – evidence from a rising credit market. *Working Paper*. <https://ideas.repec.org/p/diw/diwfin/diwfin6.3.html>
- Belleflamme, P., Omrani, N. & Peitz, M. (2015). The economics of crowdfunding platforms. *Information Economics and Policy*, 33, pp.11–28.
- Berger, S.C. & Gleisner, F. (2009) Emergence of financial intermediaries in electronic markets: the case of online P2P lending. *Business Research Journal*, 2(1), pp.39-65.
- Carr, J. H. & Megbolugbe, I. F. (1993). The federal reserve bank of boston study on mortgage lending revisited. *Journal of Housing Research*, 4(2), pp.277-313.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36(4), pp 1165-1188.
- Chen, Z., Li, B., Keung, G., Yin, H., Lin, C. & Wang, Y. (2009). How scalable could P2P live media streaming system be with the stringent time constraint? *IEEE International Conference on Communications*, pp.1-5.
- Colombo, M., Franzoni, C. & Rossi-Lamastra, C. (2015). Internal social capital and the attraction of early contributions in crowdfunding. *Entrepreneurship Theory and Practice*, 39 (1), pp.75-100.
- Culkin, N. Ekaterina M., & Andrew D. (2015). Critical innovations in the UK peer-to-peer (p2p) alternative finance markets for small firm growth. *International Journal of Entrepreneurship & Innovation*, 17(3), pp. 192-202.

- Cumming, D. & Johan, S. (2016). Crowdfunding and entrepreneurial internationalization. In: Dai, N., Siegel, D. (Eds.), *Entrepreneurial Finance: Managerial and Policy Implications*. The World Scientific Publishers.
- Cumming, D. & Johan, S. (2016). Crowdfunding and entrepreneurial internationalization. In: Dai, N., Siegel, D. (Eds.), *Entrepreneurial Finance: Managerial and Policy Implications*. The World Scientific Publishers. Chapter 5.
- Cumming, D. & Zhang, Y. (2016). Alternative investments in emerging markets: A review and new trends. *Emerging Markets Review*, 29, pp.1-23.
- Cumming, D. & Zhang, Y. (2016). Alternative investments in emerging markets: A review and new trends. *Emerging Markets Review*, 29, pp.1-23.
- Deer, L., Mi, J., Yu, Y. (2015). The rise of peer-to-peer lending in China: An overview and survey case study. ACCA publication.
- http://www.accaglobal.com/content/dam/ACCA_Global/Technical/manage/ea-china-p2p-lending.pdf . Accessed 05 May 2016.
- Diamond, D. W. (1984). Financial intermediation and delegated monitor. *Review of Economic Studies*, 51(3), pp.393-414.
- Dorflleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I. & Kammler, J. (2016). Description-text related soft information in peer-to-peer lending—evidence from two leading European platforms. *Journal of Banking & Finance*, 64(C), pp.169-187.
- Duarte, J., Siegel, S. & Young, L. (2012). Trust and credit: the role of appearance in peer-to-peer lending, *Review of Financial Studies*, 25(8), pp.2455-2484.
- Fraser, S., Bhaumik, S. & Wright, M. (2015). What do we know about entrepreneurial finance and its relationship with growth? *International Small Business Journal*, 33(1), pp.70-88.
- Fraser, S., Bhaumik, S. and Wright, M. (2015). What do we know about entrepreneurial finance and its relationship with growth? *International Small Business Journal*, 33(1),

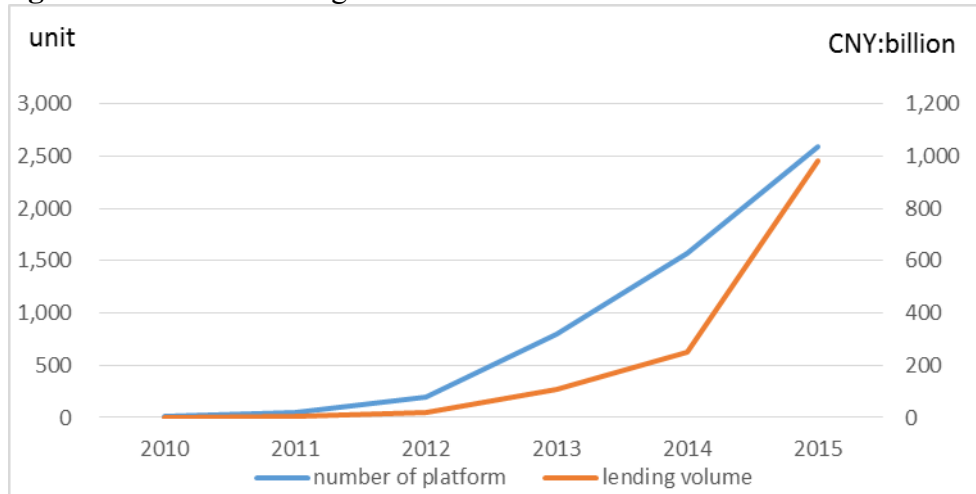
pp.70-88.

- Freedman, S. & Jin, G. Z. (2008). Do social networks solve information problems for peer-to-peer lending? evidence from prosper.com. *SSRN Electronic Journal*, pp.8-43.
- Guo, Y., Zhou, W., Luo, C., Liu, C. & Xiong, H. (2016). Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research*, 249 (2), pp. 417–426.
- Herzenstein, M. & Andrews, R. L. (2008). The democratization of personal consumer loans? determinants of success in online peer-to-peer loan auctions. *Bulletin of the University of Delaware*, 15(3), pp.274–277.
- Hota, C., Upadhyaya, S., & Al-Karaki, (2015). Advances in secure knowledge management in the big data era. *Information Systems Frontiers*, 17(5), pp983-986.
- Hulme, M. K., & Wright, C. (2006). Internet based social lending: past, present and future. *Social Futures Observatory*, (October).
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.130.3274&rep=rep1&type=pdf>.
- Iyer, R., Khwaja, A. I., Luttmer, E. F. & Shue, K. (2016). Screening peers softly: inferring the quality of small borrowers. *Management Science*, 62(6), pp.1554-1577.
- Iyer, R., Khwaja, A. I., Luttmer, E. F. P., & Shue, K. (2009). Screening in new credit markets: can individual lenders infer borrower creditworthiness in peer-to-peer lending? *Scholarly Articles*, 15242(rwp09-031).
- Kankanhalli, A., Hahn, J., Tan, S. & Gao, G. (2016). Big data and analytics in healthcare: Introduction to the special section. *Information Systems Frontiers*, 18(2), pp233-235.
- Klaft, M. (2008). Peer to peer lending: auctioning microcredits over the internet. *Social Science Electronic Publishing*.
- Klievink, B. Romijn, B., Cunningham, S. & de Bruijn H. (2016). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers* Forthcoming.

- Ladd, H. F. (1982). Equal credit opportunity: women and mortgage credit. *American Economic Review*, 72(2), pp.166-170.
- Lee, E., & Lee, B. (2012). Herding behavior in online p2p lending: an empirical investigation. *Electronic Commerce Research & Applications*, 11(5), pp.495-503.
- Lin, M., Prabhala, N., & Viswanathan, S. (2009). Social networks as signalling mechanisms: Evidence from online peer-to-peer lending. *WISE 2009*.
- Lin, M., Prabhala, R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: social networks and adverse selection in online peer-to-peer lending. *Management Science*, 59(1), pp.17-35.
- Liu, D., Brass, D. & Chen, D. (2015). Friendships in online peer-to-peer lending: pipes, prisms, and relational herding. *Mis Quarterly*, 39(3), pp.729-742.
- Loureiro, Y. & Gonzalez, L. (2015). Competition against common sense: Insights on peer-to-peer lending as a tool to allay financial exclusion. *International Journal of Bank Marketing*, 33 (5), pp.605 – 623.
- Mild, A., Waitz, M. & Wockl, J. (2015). How low can you go? — Overcoming the inability of lenders to set proper interest rates on unsecured peer-to-peer lending markets. *Journal of Business Research*, 68 (6), pp. 1291–1305.
- Poblet, M., Garcia-Cuesta, E. & Casanovas, P. (2017). Crowdsourcing roles, methods and tools for data-intensive disaster management. *Information Systems Frontiers* Forthcoming.
- Pope, D. G. & Sydnor, J. R. (2011). What's in a picture? evidence of discrimination from prosper.com. *Journal of Human Resources*, 46(1), pp.53-92.
- Ravina, E. (2008). Love & Loans: The effect of beauty and personal characteristics in credit markets. *SSRN Electronic Journal*.
- Ross, S. L. & Yinger, J. (2004). The color of credit – mortgage discrimination, research

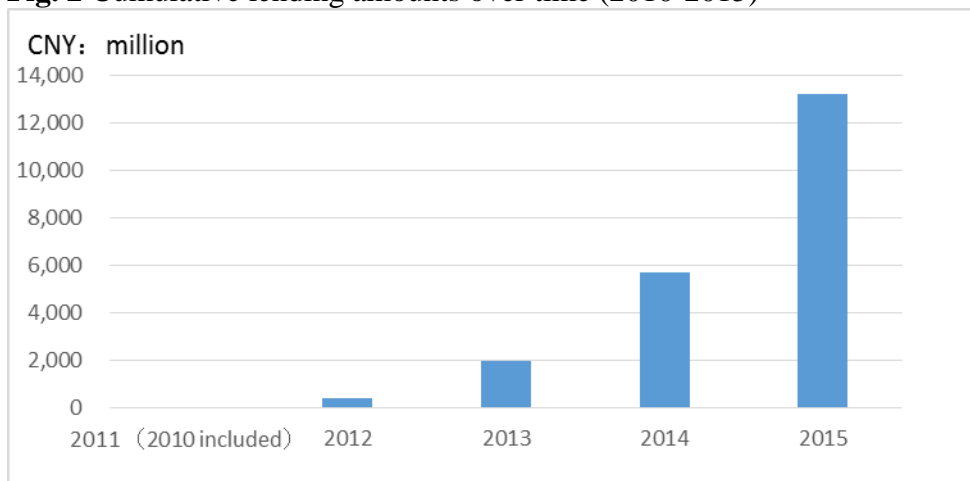
- methodology, and fair-lending enforcement. *Journal of Economics*, 81(2), pp.193-196.
- Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, 71(3), pp.393-410.
- Thuan, N.H., Antunes, P. & Johnstone (2016). Factors influencing the decision to crowdsource: A systematic literature review. *Information Systems Frontiers*, 18(1), pp47-68.
- Wang, H., Martina, G. & Jay, E. A. (2009). People-to-people lending: The emerging e-commerce transformation of a financial market. *Value creation in e-business management*, 36, pp.182-195.
- Weiss, G. N. F., Pelger, K., & Horsch, A. (2010). Mitigating adverse selection in p2p lending – empirical evidence from prosper.com. *Social Science Electronic Publishing*.
- Yuan, S. & Hsieh, C. (2016). An impactful crowdsourcing intermediary design – a case of a service imagery crowdsourcing system. *Information Systems Frontiers* Forthcoming.

Fig. 1 China's P2P lending market



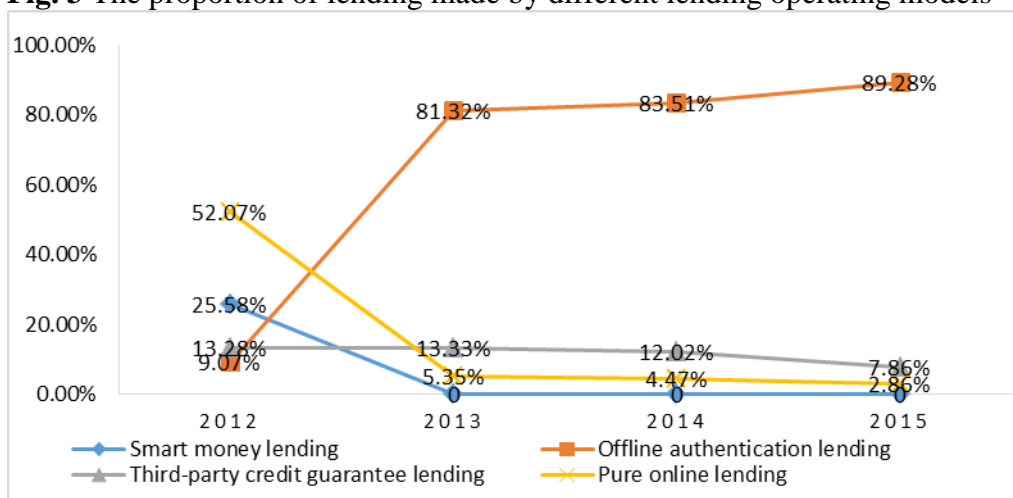
Source: WDZJ.COM

Fig. 2 Cumulative lending amounts over time (2010-2015)



Source: The annual reports of Renrendai (2011-2015)

Fig. 3 The proportion of lending made by different lending operating models



Source: The annual reports of Renrendai (2011-2015)

Table 1 Summary statistics for listings

| Variable | All listings | | | | | Offline authentication | | | Third-party guarantee | | | Pure online | | |
|---------------------------------|--------------|--------|--------|--------|--------|------------------------|--------|-------|-----------------------|--------|-------|-------------|--------|-------|
| | N | Mean | SD | Min | Max | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| Listing Outcome | | | | | | | | | | | | | | |
| LISTSUC | 197,743 | 0.404 | 0.491 | 0.000 | 1.000 | 158,438 | 0.999 | 0.034 | 21,889 | 0.987 | 0.113 | 17,416 | 0.056 | 0.231 |
| Loan Characteristics | | | | | | | | | | | | | | |
| LSIZE | 489,180 | 10.544 | 1.062 | 6.908 | 13.816 | 158,620 | 11.008 | 0.472 | 22,177 | 10.713 | 0.616 | 308,383 | 10.293 | 1.213 |
| LISTRATE | 489,180 | 0.130 | 0.025 | 0.070 | 0.240 | 158,620 | 0.121 | 0.009 | 22,177 | 0.111 | 0.010 | 308,383 | 0.135 | 0.029 |
| LLENGTH | 489,180 | 21.270 | 11.036 | 1.000 | 48.000 | 158,620 | 30.673 | 7.672 | 22,177 | 19.450 | 8.105 | 308,383 | 16.564 | 9.505 |
| Borrower Characteristics | | | | | | | | | | | | | | |
| FEMALE | 90445 | 0.185 | 0.388 | 0.000 | 1.000 | 46,097 | 0.291 | 0.454 | 5,559 | 0.251 | 0.433 | 38,789 | 0.126 | 0.332 |
| AGE | 489180 | 33.033 | 8.001 | 17.000 | 74.000 | 158,620 | 37.152 | 8.503 | 22,177 | 38.124 | 8.474 | 308,383 | 30.548 | 6.513 |
| EDU0 | 160,699 | 0.329 | 0.470 | 0.000 | 1.000 | 40,494 | 0.255 | 0.436 | 4,653 | 0.210 | 0.407 | 11,5552 | 0.375 | 0.484 |
| EDU1 | 216,576 | 0.443 | 0.497 | 0.000 | 1.000 | 80,150 | 0.505 | 0.500 | 12,232 | 0.552 | 0.497 | 12,4194 | 0.403 | 0.490 |
| EDU2 | 105,292 | 0.215 | 0.411 | 0.000 | 1.000 | 35,922 | 0.226 | 0.419 | 4,992 | 0.225 | 0.418 | 64,378 | 0.209 | 0.406 |
| EDU3 | 6,613 | 0.014 | 0.115 | 0.000 | 1.000 | 2,054 | 0.013 | 0.113 | 300 | 0.014 | 0.116 | 4,259 | 0.014 | 0.117 |
| MAR0 | 176,871 | 0.362 | 0.480 | 0.000 | 1.000 | 25,911 | 0.163 | 0.370 | 4,244 | 0.191 | 0.393 | 146,716 | 0.476 | 0.499 |
| MAR1 | 278,246 | 0.569 | 0.495 | 0.000 | 1.000 | 114,480 | 0.722 | 0.448 | 16,831 | 0.759 | 0.428 | 146,935 | 0.476 | 0.499 |
| MAR2 | 32,733 | 0.067 | 0.250 | 0.000 | 1.000 | 17,382 | 0.110 | 0.312 | 1,073 | 0.048 | 0.215 | 14,278 | 0.046 | 0.210 |
| MAR3 | 1,330 | 0.003 | 0.052 | 0.000 | 1.000 | 847 | 0.005 | 0.073 | 29 | 0.001 | 0.036 | 454 | 0.001 | 0.038 |
| Credit information | | | | | | | | | | | | | | |
| CREDA A | 316 | 0.001 | 0.025 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 316 | 0.001 | 0.032 |
| CREDA | 180,893 | 0.370 | 0.483 | 0.000 | 1.000 | 158,620 | 1.000 | 0.003 | 22,168 | 1.000 | 0.020 | 105 | 0.000 | 0.018 |
| CREDB | 503 | 0.001 | 0.032 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 503 | 0.002 | 0.040 |
| CREDC | 1,046 | 0.002 | 0.046 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 9 | 0.000 | 0.020 | 1,037 | 0.003 | 0.058 |
| CREDD | 4,111 | 0.008 | 0.091 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 4,111 | 0.013 | 0.115 |
| CREDE | 5,856 | 0.012 | 0.109 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 5,856 | 0.019 | 0.136 |
| CREDDR | 296,455 | 0.606 | 0.489 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 296,455 | 0.961 | 0.193 |
| INC0 | 1,140 | 0.002 | 0.048 | 0.000 | 1.000 | 191 | 0.001 | 0.035 | 4 | 0.000 | 0.013 | 945 | 0.003 | 0.055 |
| INC1 | 5,203 | 0.011 | 0.103 | 0.000 | 1.000 | 131 | 0.001 | 0.029 | 13 | 0.001 | 0.024 | 5,059 | 0.016 | 0.127 |
| INC2 | 172,190 | 0.352 | 0.478 | 0.000 | 1.000 | 39,145 | 0.247 | 0.431 | 4,360 | 0.197 | 0.397 | 128,685 | 0.417 | 0.493 |
| INC3 | 169,887 | 0.347 | 0.476 | 0.000 | 1.000 | 64,515 | 0.407 | 0.491 | 2,772 | 0.125 | 0.331 | 102,600 | 0.333 | 0.471 |
| INC4 | 70,651 | 0.144 | 0.352 | 0.000 | 1.000 | 24,689 | 0.156 | 0.363 | 8,852 | 0.399 | 0.490 | 37,110 | 0.120 | 0.325 |
| INC5 | 42,960 | 0.088 | 0.283 | 0.000 | 1.000 | 17,416 | 0.110 | 0.313 | 4,800 | 0.216 | 0.412 | 20,744 | 0.067 | 0.250 |
| INC6 | 27,149 | 0.055 | 0.229 | 0.000 | 1.000 | 12,533 | 0.079 | 0.270 | 1,376 | 0.062 | 0.241 | 13,240 | 0.043 | 0.203 |
| CREDITL | 489,180 | 4.020 | 5.231 | 0.000 | 13.459 | 158,620 | 9.385 | 3.979 | 22,177 | 7.252 | 5.047 | 308,383 | 1.028 | 2.988 |
| HOME | 99,183 | 0.203 | 0.402 | 0.000 | 1.000 | 94,375 | 0.288 | 0.453 | 3,318 | 0.068 | 0.252 | 126,457 | 0.169 | 0.374 |
| CAR | 224,150 | 0.458 | 0.498 | 0.000 | 1.000 | 45,701 | 0.595 | 0.491 | 1,507 | 0.150 | 0.357 | 51,975 | 0.410 | 0.492 |
| Lending Models | | | | | | | | | | | | | | |
| PURE | 308,383 | 0.630 | 0.483 | 0.000 | 1.000 | | | | | | | | | |
| OFFLINE | 158,620 | 0.325 | 0.468 | 0.000 | 1.000 | | | | | | | | | |
| GUAR | 22,177 | 0.045 | 0.208 | 0.000 | 1.000 | | | | | | | | | |

Note: Table 1 presents summary statistics for the four samples of listings. For each variable, we report the number of observations (N), the mean, and the standard deviation, as well as the minimum and maximum values. To save space, summary statistics for the control variables are omitted and available upon request.

Table 2 Summary statistics for funded loans

| Variable | All loans | | | | | Offline authentication | | | Third-party guarantee | | | Pure online | | |
|---------------------------------|-----------|--------|-------|--------|--------|------------------------|--------|-------|-----------------------|--------|-------|-------------|--------|-------|
| | N | Mean | SD | Min | Max | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| Loan outcome | | | | | | | | | | | | | | |
| LRATE | 197,743 | 0.120 | 0.011 | 0.070 | 0.240 | 158,438 | 0.121 | 0.009 | 21,889 | 0.111 | 0.010 | 17,416 | 0.124 | 0.018 |
| LDEF | 2,214 | 0.011 | 0.105 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 2,214 | 0.127 | 0.333 |
| loan Characteristics | | | | | | | | | | | | | | |
| LSIZE | 197,743 | 10.854 | 0.652 | 8.006 | 13.122 | 158,438 | 11.008 | 0.472 | 21,889 | 10.712 | 0.618 | 17,416 | 9.626 | 0.753 |
| LLENGTH | 197,743 | 27.916 | 9.666 | 3.000 | 48.000 | 158,438 | 30.674 | 7.673 | 21,889 | 19.466 | 8.125 | 17,416 | 13.453 | 8.605 |
| Borrower Characteristics | | | | | | | | | | | | | | |
| FEMALE | 53,579 | 0.271 | 0.444 | 0.000 | 1.000 | 46,031 | 0.291 | 0.454 | 5,476 | 0.250 | 0.433 | 2,072 | 0.119 | 0.324 |
| AGE | 197,743 | 36.834 | 8.480 | 20.000 | 66.000 | 158,438 | 37.155 | 8.503 | 21,889 | 38.140 | 8.477 | 17,416 | 32.269 | 6.665 |
| EDU0 | 49,105 | 0.248 | 0.432 | 0.000 | 1.000 | 40,444 | 0.255 | 0.436 | 4,582 | 0.209 | 0.407 | 4,079 | 0.234 | 0.424 |
| EDU1 | 99,332 | 0.502 | 0.500 | 0.000 | 1.000 | 80,055 | 0.505 | 0.500 | 12,088 | 0.552 | 0.497 | 7,189 | 0.413 | 0.492 |
| EDU2 | 46,457 | 0.235 | 0.424 | 0.000 | 1.000 | 35,888 | 0.227 | 0.419 | 4,926 | 0.225 | 0.418 | 5,643 | 0.324 | 0.468 |
| EDU3 | 2,849 | 0.014 | 0.119 | 0.000 | 1.000 | 2,051 | 0.013 | 0.113 | 293 | 0.013 | 0.115 | 505 | 0.029 | 0.168 |
| MAR0 | 36,619 | 0.185 | 0.388 | 0.000 | 1.000 | 25,883 | 0.163 | 0.370 | 4,175 | 0.191 | 0.393 | 6,561 | 0.377 | 0.485 |
| MAR1 | 141,096 | 0.714 | 0.452 | 0.000 | 1.000 | 114,359 | 0.722 | 0.448 | 16,616 | 0.759 | 0.428 | 10,121 | 0.581 | 0.493 |
| MAR2 | 19,142 | 0.097 | 0.296 | 0.000 | 1.000 | 17,349 | 0.110 | 0.312 | 1,069 | 0.049 | 0.216 | 724 | 0.042 | 0.200 |
| MAR3 | 886 | 0.004 | 0.067 | 0.000 | 1.000 | 847 | 0.005 | 0.073 | 29 | 0.001 | 0.036 | 10 | 0.001 | 0.024 |
| Credit Information | | | | | | | | | | | | | | |
| CREDA | 172 | 0.001 | 0.029 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 172 | 0.010 | 0.099 |
| CREDA | 180,385 | 0.912 | 0.283 | 0.000 | 1.000 | 158,438 | 1.000 | 0.000 | 21,880 | 1.000 | 0.020 | 67 | 0.004 | 0.062 |
| CREDB | 248 | 0.001 | 0.035 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 248 | 0.014 | 0.118 |
| CREDC | 476 | 0.002 | 0.049 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 9 | 0.000 | 0.020 | 467 | 0.027 | 0.162 |
| CREDD | 1,928 | 0.010 | 0.098 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 1,928 | 0.111 | 0.314 |
| CREDE | 3,233 | 0.016 | 0.127 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 3,233 | 0.186 | 0.389 |
| CREDHR | 11,301 | 0.057 | 0.232 | 0.000 | 1.000 | 0 | 0.000 | 0.000 | 0 | 0.000 | 0.000 | 11,301 | 0.649 | 0.477 |
| INC0 | 200 | 0.001 | 0.032 | 0.000 | 1.000 | 191 | 0.001 | 0.035 | 4 | 0.000 | 0.014 | 5 | 0.000 | 0.017 |
| INC1 | 175 | 0.001 | 0.030 | 0.000 | 1.000 | 131 | 0.001 | 0.029 | 13 | 0.001 | 0.024 | 31 | 0.002 | 0.042 |
| INC2 | 49,794 | 0.252 | 0.434 | 0.000 | 1.000 | 39,076 | 0.247 | 0.431 | 4,300 | 0.196 | 0.397 | 6,418 | 0.369 | 0.482 |
| INC3 | 72,726 | 0.368 | 0.482 | 0.000 | 1.000 | 64,442 | 0.407 | 0.491 | 2,692 | 0.123 | 0.328 | 5,592 | 0.321 | 0.467 |
| INC4 | 35,695 | 0.181 | 0.385 | 0.000 | 1.000 | 24,665 | 0.156 | 0.363 | 8,776 | 0.401 | 0.490 | 2,254 | 0.129 | 0.336 |
| INC5 | 23,925 | 0.121 | 0.326 | 0.000 | 1.000 | 17,407 | 0.110 | 0.313 | 4,744 | 0.217 | 0.412 | 1,774 | 0.102 | 0.302 |
| INC6 | 15,228 | 0.077 | 0.267 | 0.000 | 1.000 | 12,526 | 0.079 | 0.270 | 1,360 | 0.062 | 0.241 | 1,342 | 0.077 | 0.267 |
| SCORE | 196,704 | 5.129 | 0.304 | 0.000 | 5.501 | 158,438 | 5.193 | 0.001 | 21,889 | 5.194 | 0.007 | 16,377 | 4.425 | 0.756 |
| CREDITL | 197,743 | 9.117 | 4.065 | 0.000 | 13.459 | 158,438 | 9.384 | 3.980 | 21,889 | 7.212 | 5.062 | 17,416 | 9.091 | 2.462 |
| HOME | 51,810 | 0.262 | 0.440 | 0.000 | 1.000 | 94,261 | 0.288 | 0.453 | 3,263 | 0.068 | 0.251 | 8,938 | 0.268 | 0.443 |
| CAR | 106,462 | 0.538 | 0.499 | 0.000 | 1.000 | 45,656 | 0.595 | 0.491 | 1,482 | 0.149 | 0.356 | 4,672 | 0.513 | 0.500 |
| Lending Models | | | | | | | | | | | | | | |
| PURE | 17,416 | 0.088 | 0.283 | 0.000 | 1.000 | | | | | | | | | |
| OFFLINE | 158,438 | 0.801 | 0.399 | 0.000 | 1.000 | | | | | | | | | |
| GUAR | 21,889 | 0.111 | 0.314 | 0.000 | 1.000 | | | | | | | | | |

Note: Table 2 presents summary statistics for the four samples of loans. For each variable, we report the number of observations (N), the mean, and the standard deviation, as well as the minimum and maximum values. To save space, summary statistics for the control variables are omitted and available upon request.

Table 3 Logit regression results of the probability of funding

| | Model1 | Model2 | Model3 | Model4 | Model5 |
|-----------------------|---------------------|------------------------|------------------------|---------------------|------------------------|
| <i>CREDA</i> | 1.801*** (0.072) | | | | 1.376*** (0.074) |
| <i>CREDB</i> | 4.660*** (0.019) | | | | 1.878*** (0.130) |
| <i>CREDC</i> | 1.667*** (0.057) | | | | 1.538*** (0.059) |
| <i>CREDD</i> | 1.546*** (0.040) | | | | 1.370*** (0.042) |
| <i>CREDE</i> | 1.607*** (0.020) | | | | 1.482*** (0.022) |
| <i>CAR</i> | 1.836*** (0.017) | | | | 1.783*** (0.019) |
| <i>HOME</i> | 0.129*** (0.010) | | | | 0.170*** (0.012) |
| <i>INC1</i> | 0.079*** (0.009) | | | | 0.008 (0.010) |
| <i>INC2</i> | -0.088 (0.133) | | | | -0.049 (0.145) |
| <i>INC3</i> | 0.579*** (0.118) | | | | 0.764*** (0.130) |
| <i>INC4</i> | 0.604*** (0.118) | | | | 0.919*** (0.130) |
| <i>INC5</i> | 0.637*** (0.118) | | | | 1.106*** (0.130) |
| <i>INC6</i> | 0.754*** (0.119) | | | | 1.377*** (0.131) |
| <i>LSIZE</i> | 0.849*** (0.119) | | | | 1.647*** (0.132) |
| <i>LISTRATE</i> | | -0.233*** (0.004) | | | -0.501*** (0.005) |
| <i>LISTRATE2</i> | | 110.441*** (1.238) | | | 33.676*** (1.988) |
| <i>LLENGTH</i> | | -262.566*** (4.152) | | | -141.887*** (6.799) |
| <i>FEMALE</i> | | 0.075*** (0.000) | | | 0.005*** (0.001) |
| <i>AGE</i> | | | 0.307*** (0.007) | | 0.027** (0.013) |
| <i>EDU1</i> | | | 0.042*** (0.000) | | 0.017*** (0.001) |
| <i>EDU2</i> | | | 0.337*** (0.006) | | 0.211*** (0.011) |
| <i>EDU3</i> | | | 0.398*** (0.007) | | 0.379*** (0.012) |
| <i>MAR1</i> | | | 0.354*** (0.021) | | 0.363*** (0.032) |
| <i>MAR2</i> | | | 0.394*** (0.006) | | 0.080*** (0.010) |
| <i>MAR3</i> | | | 0.246*** (0.012) | | -0.041* (0.022) |
| <i>OFFLINE</i> | | | 0.098* (0.053) | | -0.239* (0.122) |
| <i>GUAR</i> | | | | 4.943*** (0.026) | 3.545*** (0.133) |
| <i>Other controls</i> | | | Included in estimation | | |
| No. of Obs. | 489,180 | 489,180 | 489,180 | 489,180 | 489,180 |

Notes: For brevity, some covariates are not shown. Detailed definitions of variables can be found in the Appendix. Robust standard errors are in parentheses. *, ** and *** denote that an estimate is significantly different from zero at the 10%, 5% or 1% level, respectively.

Table 4 Tobit Regression results of the interest rates on funded listings

| | MODEL1 | MODEL2 | MODEL3 | MODEL4 | MODEL5 |
|-----------------------|----------------------|----------------------|------------------------|----------------------|----------------------|
| <i>CREDA</i> | -0.008*** (0.001) | | | | -0.004*** (0.001) |
| <i>CREDA</i> | -0.004*** (0.000) | | | | -0.002** (0.001) |
| <i>CREDB</i> | 0.004*** (0.001) | | | | 0.005*** (0.001) |
| <i>CREDC</i> | 0.003*** (0.000) | | | | 0.004*** (0.000) |
| <i>CREDD</i> | -0.002*** (0.000) | | | | 0.000 (0.000) |
| <i>CREDE</i> | -0.007*** (0.000) | | | | -0.006*** (0.000) |
| <i>CAR</i> | 0.000*** (0.000) | | | | 0.001*** (0.000) |
| <i>HOME</i> | 0.002*** (0.000) | | | | -0.001*** (0.000) |
| <i>INC1</i> | -0.010*** (0.001) | | | | -0.009*** (0.001) |
| <i>INC2</i> | -0.008*** (0.001) | | | | -0.008*** (0.001) |
| <i>INC3</i> | -0.006*** (0.001) | | | | -0.006*** (0.001) |
| <i>INC4</i> | -0.008*** (0.001) | | | | -0.007*** (0.001) |
| <i>INC5</i> | -0.010*** (0.001) | | | | -0.007*** (0.001) |
| <i>INC6</i> | -0.011*** (0.001) | | | | -0.007*** (0.001) |
| <i>LSIZE</i> | | -0.004*** (0.000) | | | -0.003*** (0.000) |
| <i>LLENGTH</i> | | 0.001*** (0.000) | | | 0.001*** (0.000) |
| <i>FEMALE</i> | | | 0.000*** (0.000) | | 0.000 (0.000) |
| <i>AGE</i> | | | 0.000*** (0.000) | | 0.000*** (0.000) |
| <i>EDU1</i> | | | -0.001*** (0.000) | | -0.001*** (0.000) |
| <i>EDU2</i> | | | -0.002*** (0.000) | | -0.002*** (0.000) |
| <i>EDU3</i> | | | -0.002*** (0.000) | | -0.002*** (0.000) |
| <i>MAR1</i> | | | 0.000*** (0.000) | | -0.001*** (0.000) |
| <i>MAR2</i> | | | -0.001*** (0.000) | | -0.001*** (0.000) |
| <i>MAR3</i> | | | 0.000 (0.000) | | -0.001*** (0.000) |
| <i>OFFLINE</i> | | | | -0.003*** (0.000) | -0.010*** (0.001) |
| <i>GUAR</i> | | | | -0.017*** (0.000) | -0.018*** (0.001) |
| <i>Other controls</i> | | | Included in estimation | | |
| No. of Obs. | 197,743 | 197,743 | 197,743 | 197,743 | 197,743 |

Notes: For brevity, some covariates are not shown. Detailed definitions of variables can be found in the Appendix. Robust standard errors are in parentheses. *, ** and *** denote that an estimate is significantly different from zero at the 10%, 5% or 1% level, respectively.

Table 5 Logit Regression results of the probability of default

| | MODEL1 | MODEL2 | MODEL3 | MODEL4 |
|-----------------------|-----------------------|------------------------|----------------------|-----------------------|
| <i>CREDA</i> | 5.923 (2153.256) | | | 6.103 (2070.928) |
| <i>CREDA</i> | 0.955*** (0.250) | | | 0.649** (0.265) |
| <i>CREDB</i> | 2.036*** (0.353) | | | 1.980*** (0.360) |
| <i>CREDC</i> | 1.129*** (0.118) | | | 1.129*** (0.128) |
| <i>CREDD</i> | 1.630*** (0.094) | | | 1.599*** (0.104) |
| <i>CREDE</i> | 1.484*** (0.070) | | | 1.454*** (0.092) |
| <i>CAR</i> | 0.103*** (0.035) | | | -0.065* (0.039) |
| <i>HOME</i> | 0.077*** (0.029) | | | 0.102*** (0.034) |
| <i>INC1</i> | -5.775 (12054.840) | | | -6.180 (11249.570) |
| <i>INC2</i> | -5.547 (12054.840) | | | -5.992 (11249.570) |
| <i>INC3</i> | -5.708 (12054.840) | | | -6.046 (11249.570) |
| <i>INC4</i> | -5.728 (12054.840) | | | -6.245 (11249.570) |
| <i>INC5</i> | -5.907 (12054.840) | | | -6.464 (11249.570) |
| <i>INC6</i> | -5.996 (12054.840) | | | -6.505 (11249.570) |
| <i>LSIZE</i> | | 0.008 (0.021) | | -0.001 (0.028) |
| <i>LISTRATE</i> | | -31.438*** (5.270) | | -37.732*** (6.261) |
| <i>LISTRATE2</i> | | 75.856*** (17.575) | | 87.519*** (20.500) |
| <i>LLENGTH</i> | | -0.014*** (0.002) | | -0.008*** (0.002) |
| <i>FEMALE</i> | | | 0.084** (0.040) | 0.161*** (0.049) |
| <i>AGE</i> | | | -0.011*** (0.002) | -0.012*** (0.003) |
| <i>EDU1</i> | | | 0.192*** (0.030) | 0.151*** (0.036) |
| <i>EDU2</i> | | | 0.595*** (0.036) | 0.438*** (0.043) |
| <i>EDU3</i> | | | 0.873*** (0.104) | 0.680*** (0.130) |
| <i>MAR1</i> | | | 0.002 (0.031) | 0.017 (0.038) |
| <i>MAR2</i> | | | -0.290*** (0.062) | -0.173** (0.072) |
| <i>MAR3</i> | | | -0.320 (0.435) | -0.173 (0.518) |
| <i>Other controls</i> | | Included in estimation | | |
| Log likelihood | -5624.970 | -6249.650 | -6258.170 | -4440.060 |
| No. of Obs. | 17,416 | 17,416 | 17,416 | 17,416 |

Notes: For brevity, some covariates are not shown. Detailed definitions of variables can be found in the Appendix. Robust standard errors are in parentheses. *, ** and *** denote that an estimate is significantly different from zero at the 10%, 5% or 1% level, respectively.

Table 6 Regression results for online and offline sub-samples

| | Funding success | | Interest rate | |
|-----------------------|--------------------------|------------------------|------------------------|----------------------|
| | OFFLINE | ONLINE | OFFLINE | ONLINE |
| <i>CREDA</i> | | 1.372*** (0.074) | | -0.004*** (0.001) |
| <i>CREDA</i> | | 1.847*** (0.132) | | -0.004* (0.002) |
| <i>CREDB</i> | | 1.535*** (0.060) | | 0.004*** (0.001) |
| <i>CREDC</i> | | 1.362*** (0.042) | | 0.003*** (0.001) |
| <i>CREDD</i> | | 1.478*** (0.022) | | 0.000 (0.000) |
| <i>CREDE</i> | | 1.780*** (0.019) | | -0.005*** (0.000) |
| <i>CAR</i> | 0.060 (0.056) | 0.183*** (0.012) | 0.000*** (0.000) | 0.001*** (0.000) |
| <i>HOME</i> | -0.083 (0.057) | 0.007 (0.010) | 0.000 (0.000) | -0.001*** (0.000) |
| <i>INC1</i> | 0.257 (1763.996) | 0.061 (0.176) | -0.011*** (0.000) | -0.028*** (0.008) |
| <i>INC2</i> | -2.995 (1117.072) | 0.903*** (0.162) | -0.008*** (0.000) | -0.029*** (0.007) |
| <i>INC3</i> | -2.913 (1117.072) | 1.069*** (0.162) | -0.005*** (0.000) | -0.029*** (0.007) |
| <i>INC4</i> | -2.959 (1117.072) | 1.244*** (0.162) | -0.005*** (0.000) | -0.030*** (0.007) |
| <i>INC5</i> | -2.843 (1117.072) | 1.561*** (0.163) | -0.006*** (0.000) | -0.030*** (0.007) |
| <i>INC6</i> | -2.937 (1117.072) | 1.841*** (0.164) | -0.007*** (0.000) | -0.029*** (0.007) |
| <i>LSIZE</i> | 0.110* (0.067) | -0.516*** (0.006) | -0.004*** (0.000) | -0.001*** (0.000) |
| <i>LISTRATE</i> | -899.165*** (131.008) | 29.166*** (2.023) | | |
| <i>LISTRATE2</i> | 3782.664*** (548.064) | -127.501*** (6.884) | | |
| <i>LLENGTH</i> | 0.003 (0.005) | 0.006*** (0.001) | 0.001*** (0.000) | 0.001*** (0.000) |
| <i>FEMALE</i> | -0.072 (0.051) | 0.038*** (0.014) | 0.000*** (0.000) | -0.001*** (0.000) |
| <i>AGE</i> | 0.013*** (0.003) | 0.017*** (0.001) | 0.000*** (0.000) | 0.000*** (0.000) |
| <i>EDU1</i> | 0.059 (0.056) | 0.217*** (0.011) | 0.000*** (0.000) | 0.000 (0.000) |
| <i>EDU2</i> | 0.097 (0.072) | 0.392*** (0.013) | -0.001*** (0.000) | -0.002*** (0.000) |
| <i>EDU3</i> | -0.073 (0.198) | 0.385*** (0.033) | -0.001*** (0.000) | -0.003*** (0.001) |
| <i>MAR1</i> | -0.124* (0.072) | 0.078*** (0.011) | -0.001*** (0.000) | 0.000 (0.000) |
| <i>MAR2</i> | -0.325*** (0.093) | -0.047** (0.024) | -0.001*** (0.000) | -0.001** (0.001) |
| <i>MAR3</i> | 3.220 (557.924) | -0.385** (0.154) | -0.001*** (0.000) | -0.007 (0.005) |
| <i>Other controls</i> | | | Included in estimation | |
| No. of Obs. | 158,620 | 308,383 | 158,438 | 17,416 |

Notes: For brevity, some covariates are not shown. Detailed definitions of variables can be found in the Appendix. Robust standard errors are in parentheses. *, ** and *** denote that an estimate is significantly different from zero at the 10%, 5% or 1% level, respectively

Appendix: Variables and Definitions

| Variable name | Description |
|---|---|
| Outcomes | |
| Funding success (<i>LISTSUC</i>) | A dummy variable, 1 if the listing was fully funded and approved; 0 otherwise. |
| Loan's interest rate (<i>LRATE</i>) | The interest rate on funded loan. |
| Loan default (<i>LDEF</i>) | A dummy variable, 1 if the successful listing defaults; 0 otherwise. |
| Loan Characteristics | |
| Loan size (<i>LSIZE</i>) | Logarithm of loan amount requested by the borrower. |
| Loan length (<i>LLENGTH</i>) | The number of months the borrower would like to pay off the loan. |
| Listing's interest rate (<i>LISTRATE</i> & <i>LISTRATE2</i>) | The interest rate specified by the borrower at time of listing and its quadratic term. |
| Personal Characteristics | |
| Gender (<i>FEMALE</i>) | A dummy variable, 1 if the borrower is female; 0 otherwise. |
| Age (<i>AGE</i>) | The age of the borrower. |
| Education dummies (<i>EDUn</i>) | A series of dummy variables indicating the education level that the borrower specified, where $n = 0$ (high school or lower, this is the baseline and not included in regressions), 1 (junior college), 2 (undergraduate), and 3 (Master's degree or above). |
| Marital status dummies (<i>MARn</i>) | A series of dummy variables indicating the marital status that the borrower specified, where $n = 0$ (unmarried; this is the baseline and not included in regressions), 1 (married), 2 (divorced) and 3 (widow). |
| Listing types | |
| Pure online listings (<i>ONLINE</i>) | This is the baseline and is not included in regressions. |
| Offline authentication listings (<i>OFFLINE</i>) | A dummy variable, 1 if the borrower's information has been verified through the offline authentication process; 0 otherwise. |
| Referrals from third-party credit guarantee companies (<i>GUAR</i>) | A dummy variable, 1 if the listing is an offline referral from third-party credit companies; 0 otherwise. |
| Hard financial information | |
| Credit grade dummies (<i>CREdn</i>) | A series of dummy variables indicating the borrower's credit grade, where $n = AA$ (credit grade is AA), A (credit grade is A), B (credit grade is B), C (credit grade is C), D (credit grade is D), E (credit grade is E), and HR (credit grade is HR; this is the baseline and not included in regressions). |
| Personal monthly income dummies (<i>INCn</i>) | A series of dummy variables indicating the monthly income range that the borrower specified, where $n = 0$ (less than 1,000 RMB, this is the baseline and not included in regressions), 1 (between 1,000 and 2,000 RMB), 2 (between 2,001 and 5,000 RMB), 3 (between 5,001 and 10,000 RMB), 4 (between 10,001 and 20,000 RMB), 5 (between 20,001 and 50,000 RMB) and 6 (over 50,000 RMB). |
| Home owner (<i>HOME</i>) | A dummy variable 1, if the borrower owns a house; 0 otherwise. |
| Car owner (<i>CAR</i>) | A dummy variable 1 if the borrower owns a car; 0 otherwise. |
| Other control variables | |
| Loan category dummies (<i>LCn</i>) | A series of dummy variables indicating the reason for borrowing that the borrower specified, where $n = 0$ (consolidate debt; this is the baseline and not included in regressions), 1 (buying a house), 2 (bill payments), 3 (business loans), 4 (automobile loans), 5 (home improvement), 6 (medical loans), 7 (education loans), 8 (wedding loans), and 9 (other loans). |
| Length of loan description (<i>LTEXT</i>) | Logarithm of the total length of text on the borrower's listing. |
| Employment status (<i>EMPn</i>) | A series of dummy variables indicating the employment status that the borrower specified, where $n = 0$ (runs online business; this is the baseline and not included in regressions), 1 (wage/salary earner), 2 (self-employed) and 3 (others). |
| Length of employment status dummies (<i>EMPYEAR</i>) | A series of dummy variables indicating the range of borrower employment length in years, where $n = 0$ (less than 1 year; this is the baseline and not included in regressions), 1 (between 1 and 3 years), 2 (between 4 and 5 years), and 3 (over 5 years). |
| Geographical region dummies (<i>REGn</i>) | A series of dummy variables indicating the region the borrower comes from, where $n = 0$ (Centre and south; this is the baseline and not included in regressions), 1 (east coast), 2 (north), 3 (southwest), 4 (northeast), 5 (northwest), and 6 (municipalities). |

