

GENDER EFFECT ON THE DEFAULT RISK IN PEER-TO-PEER LENDING MARKETS: THE CASE OF THE LARGEST CHINESE PLATFORM

Lin Lingnan *

* Tsinghua University, China

Contact details: Institute of Economics, School of Social Sciences, Tsinghua University, 100084, Beijing, China



Abstract

How to cite this paper:

Lingnan, L. (2019). Gender effect on the default risk in peer-to-peer lending markets: The case of the largest Chinese platform. *Risk Governance and Control: Financial Markets & Institutions*, 9(3), 8-22.
<http://doi.org/10.22495/rgcv9i3p1>

Copyright © 2019 The Authors

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>

ISSN Online: 2077-4303

ISSN Print: 2077-429X

Received: 14.03.2019

Accepted: 16.05.2019

JEL Classification: E44, G21, J16

DOI: 10.22495/rgcv9i3p1

Research of gender effect on funding success in peer-to-peer lending markets demonstrates that gender discrimination is a platform-specific phenomenon rather than a common feature. Can we get a similar conclusion about the relationship between gender and credit risk? How do gender differences affect default risk? We try to answer this question using the data of the largest peer-to-peer lending platform RenRenDai spanning from March 2016 to September 2016. In order to avoid the endogeneity problem, this paper first uses the instrumental variable method to conduct a baseline Probit model estimate connecting gender difference to the default rate with several borrowers' individual characteristics under control. Then the original Probit model and a propensity score matching method aiming to eliminate the effects of divergent observable characteristics are applied to test the robustness of the outcome. Both the baseline estimation and the robustness test show that there is no significant gender effect on the probability of default, *ceteris paribus*. Therefore, borrowers' gender is not a good screening device for the P2P lending platform to control the credit risk; other factors should be taken into account to reduce the non-performing loan rate. However, since this paper only investigates the situation of RenRenDai and the data we use is limited, we should be very careful to generalize our findings to other P2P lending platforms. More research on different P2P lending platforms in different regulatory regimes is in necessity.

Keywords: Peer-to-Peer Lending, Gender Effect, Default Risk

Authors' individual contribution: the author is responsible for all the contributions to the paper according to CRediT (Contributor Roles Taxonomy) standards.

1. INTRODUCTION

Online Peer-to-Peer (P2P) lending is an emerging internet financial mode with the advent of Web 2.0 technology. In the online P2P market, investors who have a surplus of funds lend money to applicants for loans on the internet, instead of through conventional financial institutions such as banks. In the absence of conventional financial institutions, online P2P lending renders a higher return to lenders and lower interest rate to borrowers by lowering transaction cost. With so many advantages, online P2P lending has exploded in recent years. The

first online P2P platform, Zopa, was born in March 2005 in the United Kingdom. Thereafter, online P2P platforms represented by Prosper, Lending Club of United States, sprang up in many countries. Because of the maturity of the credit report system, online P2P lending is growing rapidly in western countries. Online P2P lending commenced in China in 2007. Accompanied by interest rate deregulation, finance disintermediation and the burst of informal lending, online P2P lending has grown violently and an abundance of online P2P platforms has emerged since 2010.

The goal of this paper is to estimate the gender effect on default risk in the online P2P lending markets. Using data from the largest Chinese lending platform RenRenDai, we test the effect of borrowers' gender on the probability of default based on a multivariate Probit model regression. The test results show that there is no significant gender effect on the probability of default. Males and females seem to have equal default risk in P2P lending, *ceteris paribus*. This finding remains true in a variety of robustness check. Therefore, we can be confident enough to confirm that no significant gender effect exists on RenRenDai.

Compared to existing research, this paper has several distinguishing features. First, our study focuses on the gender effect on the default risk on P2P lending platforms, which is limited in the existing literature. Second, compared to the analysis based on the data from Prosper, Lending Club and other platforms in different countries, our data is collected by some internet technology which can simultaneously procure the information submitted by loan applicants. At RenRenDai, applicants should submit a bunch of information for verification, including gender, age, monthly salaries, education level, marital status, employment status, loan purposes, etc. The submission of personal information is mandatory and far more substantial than other P2P lending platforms such as Prosper, Lending Club. The abundance of information about individual characteristics allows us to take more influencing factors under control, which enables more accurate estimation about the gender effect on default risk and makes our results more precise.

The remainder of this paper is organized as follow. Section 2 takes a review of the existing literature in the field of P2P lending. Section 3 introduces our test methodology and regression model briefly. Section 4 takes an overview of lending and borrowing on RenRenDai and describes the data we used. Section 5 describes the econometric results employing a two-step instrumental Probit regression. Then, we take a few robustness checks to confirm the adaptability of our test results in Section 6. In the last section, we get the final conclusion and point out several reasons why generalization of this conclusion needs further research.

2. LITERATURE REVIEW

Peer-to-Peer lending has been gaining growing popularity and more and more research attraction in this area. Hulme and Wright (2006) studied the case of Zopa and concluded that the emergence of P2P lending reflects the social trends and a need for reform of the financial system in the information era. Ashta and Assadi (2008) examined the role that internet techniques play in promoting social interactions and associations with lower cost for P2P lending. De Roure et al. (2018) developed a simple theoretical model to test how peer-to-peer (P2P) platforms compete with banks for loans and predicted that (i) P2P lending grows when some banks are faced with exogenously higher regulatory costs; (ii) P2P loans are riskier than bank loans; and (iii) the risk-adjusted interest rates on P2P loans are lower than those on bank loans.

A great number of researchers investigated the determinants of default risk. Gomez and Santor

(2003) found that default rates of peer group lending are lower than that of conventional individual lending in Canada. Iyer et al. (2009) found that lenders are able to use available information to infer one-third of credit risk captured by borrowers' credit scores in P2P lending. Lin et al. (2013) analyzed the social connections in determining default rate and found that borrowers' friendship with lenders is associated with lower ex-post default rates at Prosper, one of the largest P2P lending platform in the United States. Polena and Regner (2018) studied the determinants of borrowers' default in P2P lending with a new data set consisting of 70,673 loan observations from the Lending Club. They discovered that the debt-to-income ratio, inquiries in the past six months and a loan intended for a small business are positively correlated with the default rate. Annual income and credit card as loan purpose are negatively correlated.

There is also extant literature related to the role of fundamental borrower characteristics and personal information in determining default risk. Barasinska and Schafer (2010) studied the gender effect of applicants on the funding success at Smava, the largest P2P lending platform in German, and found that there is no gender discrimination in P2P lending market. Using the data from POSper, Herzenstein et al. (2008) investigated the influence of demographic attributes, including gender, race and marital status, on the funding success. Pope and Sydnor (2011) focused on the impact of gender on the funding success and interest rate and found that there is a systematic underestimate in the default rate of African Americans. Duarte et al. (2012) discovered that borrowers whose appearances are more trustworthy would have higher credit scores, more funding success and lower default risk on P2P lending platforms. Freedman and Jin (2017) examined whether social networks facilitate online markets using data from POSper and found that borrowers with social ties are consistently more likely to have their loans funded and receive lower interest rates. However, most borrowers with social ties are more likely to pay late or default.

Research that focuses on the default risk at Chinese P2P lending platform is limited. Liao Li et al. (2015) studied the case of RenRenDai and found that higher education level is associated with higher self-constraint and lower default rate, but there exists a bias in identifying credit risk through education level for lenders. Therefore, our study is expected to shed some lights on the relationship between gender and credit risk. Using data from a large P2P platform in China, Xuchen Lin et al. (2017) explored the factors that determine the default risk based on the demographic characteristics of borrowers and proposed a credit risk evaluation model which can quantify the default risk of each P2P loan. Zhang and Chen (2017) used in RenRendai to examine the dynamic relationship between prior cumulative bids and current bids. They found that lenders appear to imitate each other's behavior and herd in the P2P lending market.

3. RESEARCH METHODOLOGY

As we mentioned before, credit risk is the main concern of the research on the P2P platforms. The credit risk can be depicted by the default rate of loans. For a borrower on a P2P platform, we can

index him/her with i . A loan related to borrower i can be depicted by three basic elements: nominal interest rate I_i , duration D_i and loan amount L_i . Apparently, the borrower's probability of default p_i is correlated with nominal interest rate I_i , duration D_i and loan amount L_i . Lenders on the platform cannot observe p_i . However, they can derive p_i from observable variables related to the borrower's characteristics. We can use vector X_i to capture all the observable variables deemed to be necessary to enter the determinants of p_i .

The funding success of lending on P2P platforms is determined by the willingness of lenders to provide funds and how much lenders would provide. Whether or not lenders provide funds and the number of funds actually provided is determined by the expected return from lending. Obviously, the expected return is closely related to nominal interest rate I_i , duration D_i , loan amount L_i and default rate p_i .

Our research problem is that whether male and female borrowers have different default rate p_i given that the loan term and all the observable characteristics are the same. If this can be

$$\text{Prob}(\text{default}_i = 1) = \phi(\mathbf{X}, \boldsymbol{\beta}) = \phi(\beta_0 + \beta_1 \text{Male}_i + \beta_2 I_i + \beta_3 D_i + \beta_4 L_i + \mathbf{X}'_i \boldsymbol{\beta}_5) \quad (1)$$

where $\phi(\cdot)$ is the standard normal cumulative distribution function, and $\mathbf{X}' = (1, \text{Male}_i, I_i, D_i, L_i, \mathbf{X}'_i)$, $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \boldsymbol{\beta}'_5)$. \mathbf{X}_i is a vector of variables capturing all the observable characteristics of borrowers, and $\boldsymbol{\beta}_5$ is a coefficient vector. Hence, the probability of not default can be expressed as:

$$\text{Prob}(\text{default}_i = 0) = 1 - \phi(\mathbf{X}, \boldsymbol{\beta}) \quad (2)$$

Using Eq. 1 and Eq. 2, we can write the likelihood function as follows:

$$\mathcal{L} = \prod_{i=1}^n [\phi(\mathbf{X}, \boldsymbol{\beta})]^{\mathbf{1}(\text{default}_i=1)} [1 - \phi(\mathbf{X}, \boldsymbol{\beta})]^{1-\mathbf{1}(\text{default}_i=1)} \quad (3)$$

where $\mathbf{1}(\cdot)$ is an indicative function signaling whether or not the borrower is default.

Hence, our task is to solve the maximization problem of the log-likelihood function:

$$\max_{\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \boldsymbol{\beta}_5} \ln \mathcal{L} = \sum_{i=1}^n \{ \mathbf{1}(\text{default}_i = 1) \ln \phi(\mathbf{X}, \boldsymbol{\beta}) + [1 - \mathbf{1}(\text{default}_i = 1)] \ln [1 - \phi(\mathbf{X}, \boldsymbol{\beta})] \} \quad (4)$$

The coefficients $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \boldsymbol{\beta}_5$ are estimated through the above maximization program. The estimators can be written as $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\boldsymbol{\beta}}_5$.

The variable of interest is the dummy variable of Male. If the borrower is male, $\text{Male}=1$, else $\text{Male}=0$. The effect of gender on the probability of default is reflected by the coefficient estimator $\hat{\beta}_1$. More specifically, $\hat{\beta}_1 > 0$ ($\hat{\beta}_1 < 0$) represents that males (females) have larger default rate, respectively.

4. DATA

4.1. An overview of RenRenDai

RenRenDai, short for RenRenDai Business Consultant (Beijing) Ltd., was founded in May 2010. RenRenDai is affiliated with Youxin Financial Co. and is an independent brand. It is one of the earliest founded and leading P2P platforms, and is on the list of Top 100 Chinese Internet Enterprises in 2015 and 2016. RenRenDai aims at provide high-quality, professional financial information service for clients, build a trustworthy investment and credit lending platform.

authenticated, then let the default rate of male borrowers be p , the default rate of female borrowers would be $p + \delta$, with $\delta \neq 0$. Hence, the profit maximizing lender would have the incentive to charge a higher risk premium from a borrower with higher default rate, which can affect the probability of funding success prominently. Lenders can also use this phenomenon as a screen device to phase out borrowers of a particular gender, which constitutes another problem on the P2P markets — gender discrimination discussed in many literatures related to this area. Therefore, we will test the hypothesis below:

H1: Assumed that borrowers of different sex face the same loan terms and have similar observable characteristics, the probability of default would be different across gender groups, *ceteris paribus*.

The test of this hypothesis in the remainder of this paper will be based on the most popular Discrete Response Model - Probit Model. According to this model, the probability of default for borrower i relies on the following equation:

According to the 2018 annual report disclosed on the internet, a sum of 409,592 loans in a total volume of ¥ 30,191,885,600 had been transacted on RenRenDai, with the year-on-year growth rate of 38.47% and 49.34%, respectively. At the end of 2018, there are 18,149,168 registered users on the platform in the aggregate; the number of lenders and borrowers is 519,010 and 394,617, respectively. The number of transactions and the total amount increased continuously (see Figure 1 in Appendix).

Loan applications. In order to issue a loan application on RenRenDai, borrowers have to pass two verifications — registration and loan application verifications. First, they need to register on the platform, and submit the personal information, including education level (see Table 1 in Appendix), monthly salary, marital status, whether or not they have outstanding debts, whether or not they own a house/car and other required materials. The platform will evaluate based on the submitted information and assign a credit line/score to every applicant. After the first verification, borrowers could submit their loan applications. The platform

would investigate the authenticity of applications, and decide whether or not to issue applications.

Credit ratings. Due to the imperfect credit system and fragmentation of credit information in China, RenRenDai combines the on-line and off-line verification in the purpose of controlling loan risks. The on-line verification is based on the data analysis of the submitted applicants' information. The off-line verification is conducted by employees from cooperating credit agencies through field investigation. RenRenDai will evaluate applicants' creditworthiness through the individual credit risk analysis system with on-line and off-line verifications taken together. Then, RenRenDai assigns a final credit line and rating to each applicant. The borrowers' ratings are divided into 7 categories: AA, A, B, C, D, E, HR.

Risk reserve. In case of loan delinquencies, RenRenDai has set up a risk reserve account. When a loan applicant gets funded, a fraction of loan is charged by RenRenDai based on the applicant's rating (see Table 2). This fund is deposited in the risk reserve account. When the payment of a loan is overdue for 30 days, RenRenDai will employ the risk reserve fund to buy lenders' claims and ensure that lenders' principals are entirely paid. Due to the change of regulatory policies, RenRenDai is turning to third-party factoring mode. In this pattern, if a borrower remains in arrears for 30 days, the claims of lenders are sold to a collection or a commercial factoring agency. After buying the claims, agencies get the legal right to recover debts by all means at their disposal.

4.2. Data set

Our data set incorporates all the applications posted at the RenRenDai from March 2016 to September 2016 which received at least one bid from lenders. These data are collected using a crawler search technology, and all these data concerning loan applicants or borrowers are observable to lenders on the platform and researchers, which would not invoke any proprietary problems.

According to the data set, a total of 81,223 individuals applied for loans and received some funds larger than 0. As we can see from Table 3, females accounted for 26,091 (32.12%) and males accounted for 55,132 (67.88%) of all borrowers. The distributions of monthly salary and the definitions of variables are shown in Table 4 and Table 5, respectively. Descriptive statistics of the variables are given in Table 6. As we can see from Table 7, there exist some discrepancies between borrowers of different genders. In general, females procured larger loan amount and longer loan duration than males and paid 0.015 percent lower interest rate than males. There are also many gender differences in borrowers' personal characteristics. For example, the number of males who have a high school, undergraduate, graduate degree is larger than that of females, but more females have a college degree. The majority of borrowers have ratings A, however, an obvious increase of the number of males can be detected as the rating category degrade from B to HR. Similarly, male borrowers earn, on average, more than females, and are elder than female borrowers. The gender differences of most employment statuses are significant, except that the number of

females who are sole proprietors is not so much different from that of males compared to other occupations.

Figure 2 plots the distribution of borrowers by loan purpose. Generally speaking, liquidity consolidation and personal consumption account for most proportions of loan purpose despite gender differences. There is also some evidence in accordance with popular stereotypes: males prevail in the categories related to a car purchase, house purchase, marriage preparation, education & training, furnishing, while females predominate in other areas. However, counterintuitively, more females seem to borrow for investing in innovation activities than males.

5. ECONOMETRIC RESULTS

We can use the methodology introduced in Section 3 to conduct a baseline regression. However, it is obvious that borrowers can offer higher loan rates and lower loan amount to increase the chance of funding success. The less creditworthy borrowers are, the more incentive they have to offer tempting loan terms to be fully funded, which results in a higher probability of default. Hence, the interest rate and loan amount would be endogenous in our model. The endogeneity problem is widely discussed in the statistical literature (Heckman, 1978). The endogeneity of factors will cause the estimation model aforementioned in this paper to produce biased estimates. This bias can be remedied through a two-step regression method whereby endogenous variables are substituted by a few instrumental variables.

In the first step, two subsidiary OLS regressions are employed. The interest rate and loan amount are regressed with respect to a set of exogenous variables including loan duration, ratings, education levels, monthly salary, employment status, age, length of description, marital status, ownership of houses, ownership of autos, loan purpose. The estimation results of these two subsidiary regressions are shown in Panel A of Table 8. Most margin effects of exogenous variables on the interest rate and loan amount are significant except for gender, although the adjusted- R^2 is only 0.446 and 0.232, respectively.

After the first-step regression, the fitted value of interest rate and loan amount can be used as instrumental variables in the second-step probit model. In order to satisfy the just-identified order condition of instrumental variable, some of the exogenous variables in the first-step regression should be removed out of the second step. The employment status is selected in the removal. The choice is based on the fact that employment status is obviously a factor influencing loan terms. For instance, a borrower employed by a private enterprise would be more unpredictable in payment of loan than a civil servant due to the attributes of the job. Thus, a higher interest rate and a lower loan amount are requested for a borrower working in a private enterprise. However, given that the credit risk caused by different employment status is compensated by loan terms, the probability of default should be similar.

The estimation result is recorded in Panel B of Table 8. The bottom row of Panel B is a Wald test for

the exogeneity of interest rate and loan amount. The test statistic is insignificant, which means we cannot reject the null hypothesis of exogeneity and the endogeneity problem of interest rate and loan amount can be neglected. Furthermore, the estimated coefficient of variable Male is not significant in the two-step regression, which means that gender difference has barely an effect on the probability of default.

6. ROBUSTNESS CHECK

6.1. The number of control variables

The above two-step instrumental Probit regression shows that the problem of endogeneity is not serious. Therefore, we can use the Probit model to test if the result that gender effect on the default risk is not significant holds for the normal regression.

Table 9 reports the coefficients of independent variables determining the probability of default. Column (1) returns the outcomes of the baseline regression which incorporates a dummy variable Male, a set of variables related to loan terms and a set of dummy variables generated by borrowers' ratings. Column (2) returns the result of regression with monthly salary, age, description, ownership of houses, ownership of cars included besides the variables in the baseline regression. Column (3), (4), (5) and (6) describes the results of extended regressions with education level, employment status, marital status, loan purpose incorporated, respectively.

Each regression predicts a weak link between loan term variables — the interest, duration, loan amount and the probability of default in accordance with the results in Section 5. It is very exceptional compared with the situation in regular financial institutions. Usually, lenders will associate longer duration and higher amount with more uncertainty in loan repayment, hence higher risk premia will be required by lenders which give rise to a higher probability of default by borrowers. Hence, we testify that online P2P lending is very distinct from the conventional financial market.

It is not surprising that each model shows a strong connection between borrowers' ratings and the default rate. The probability of default increases by the degrading of rating from A to HR. Likewise, some borrower features are reported as playing a key role in the default rate. For example, the default rate decreases prominently with the growing of borrowers' education level. In comparison with borrowers who do not report their marital status, borrowers who are divorced, married and unmarried seem to have a higher default rate. This is a little bit complicated to comprehend. Yet, some individuals who borrow on the internet may be more concerned about their own privacy and refuse to submit their marital status to the P2P platforms. These borrowers are likely to attach more importance to their creditworthiness so that they have a lower probability of default. Moreover, we find a positive relationship between age and default rate. However, monthly salary, ownership of houses, ownership of cars, employment status and loan purpose appear to be irrelevant to the probability of default. It justifies our choice of employment status as the instrumental

variable to interest rate and loan amount in Section 5 again.

It is noted that in all specifications, the sign of the estimated coefficient of gender is positive. However, each regression predicts that gender has no significant effect on the probability of default. Even though we add the control variables in the regression step by step, gender remains insignificant in determining default rate.

6.2. Effects of divergent observable characteristic variables

The estimation result from the first step regression shows that males procure higher interest rate and less loan amount than females (Panel A of Table 8). Apart from the interest rate and loan amount, there are significant differences between observable characteristic variables as Table 7 shows. Plenty of divergences related to observable characteristics between different gender groups will render our *ceteris paribus* assumption invalid, and the estimation of gender effect will be inconsistent. To solve this sample problem, we use the method of propensity score matching designed for dealing with treatment effect (Rosenbaum and Rubin, 1985). The goal is to estimate the gender effect with a sample of matched individuals. Except for the gender divergence, all the other observable characteristics of these individuals are almost the same.

The resemblance of borrowers is based on their propensity score. A propensity score represents the possibility that a borrower is male given the observable characteristics and loan terms, and is calculated with a logit regression model whereby a dummy variable Male is regressed on all the other observable variables. The balance of samples in different gender groups indicated by standardized bias is shown in Figure 3. We can see that most dissimilarities of observable variables (including loan term variables) are reduced remarkably after matching. The distribution of propensity scores is plotted in Figure 4. Obviously, the distribution of propensity scores by different groups is close to each other. Indeed, only 20 males and 19 females fall outside the common support of propensity score, which means approximately 0.05% of the overall samples remain unmatched and there is a good chance that we can find an identical "twins" of the opposite gender for every borrower.

After matching with propensity score, we estimate the average gender effect on the probability of default using the kernel matching method (Heckman et al., 1998) with Epanechnikov kernel to assign a weight to every matched sample. The 39 samples falling outside of the common support are excluded from this calculation. At the same time, the bootstrap method is employed to estimate the standard error and p-value of gender effect. The result shows that the average gender effect is equal to 0.0007 and is statistically insignificant. Hence, the robustness check again justifies our result obtained earlier.

7. CONCLUSIONS

The above analysis based on the samples from the largest P2P platform RenRenDai does not inform us of any significant gender effect on the probability of default with all the observable characteristics and

loan terms being controlled. The result obtained is subjected to different robustness checks. Even if we take the endogeneity problem of the interest rate and loan amount into an account and separate out the effects of observable characteristics, this result holds true. Therefore, we can argue that there is no gender difference in the default rate on RenRenDai at this moment.

However, we must be very careful to generalize this conclusion to other P2P lending platforms due to the following reasons for two respects. First is caused by the limitation of our data set. The time span of our data is only two quarters due to the availability. However, people usually need to borrow money for smoothing their own consumption and financing other expenditures in the 1st and 4th quarters in the context of the spring festival. Fewer loans would take place in the 2nd and 3rd quarters since most firms would conduct the bonus payment to employees at that time. Therefore, the seasonal factors should be taken into account when we discuss the problem concerning default risks. Furthermore, most of the loans we use for analysis

have not yet matured. The probability of default remains uncertain on maturity. The repayment state at maturity should be investigated and the default at maturity should be distinguished from that in fulfillment. In this sense, panel data are called for to test the validity of our conclusion regarding gender effects on credit risks.

Another reason to be cautious of the generalization stems from the sampling limitation of this paper. Apparently, results obtained from different P2P lending platforms rely on the specific operational mechanism of platforms, and even the regulatory environment the specific platform dwells in. Most of the current research, including this paper, is based on the data of one single platform sample. It would be unsound to conclude that there is no gender effect on default risk on all P2P lending platforms. Therefore, a comparative analysis of divergent P2P lending platforms regarding the role of different operational mechanisms and regulatory environments is necessary and remains to be carried out in further studies.

REFERENCES

- Ashta, A., & Assadi, D. (2009). Do social cause and social technology meet? Impact of web 2.0 technologies on peer-to-peer lending transactions. *Cahiers du CEREN*, 29, 177-192. Retrieved from <https://ssrn.com/abstract=1281373>
- Barasinska, N., & Schäfer, D. (2010). *Does gender affect funding success at the peer-to-peer credit markets? Evidence from the largest German lending platform* (Working Paper). <https://doi.org/10.2139/ssrn.1738837>
- De Roure, C., Pelizzon, L., Anjan, V., & Thakor, A. V. (2018). *P2P Lenders versus banks: Cream skimming or bottom fishing?* (SAFE Working Paper No. 206). Retrieved from <https://www.econstor.eu/bitstream/10419/178504/1/1020789387.pdf>
- Duarte, J., Siegel, S., & Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies*, 25(8), 2455-2484. <https://doi.org/10.1093/rfs/hhs071>
- Freedman, S., & Jin, G. Z. (2017). The information value of online social networks: Lessons from peer-to-peer lending. *International Journal of Industrial Organization*, 51, 185-222. <https://doi.org/10.1016/j.ijindorg.2016.09.002>
- Gomez, R., & Santor, E. (2003). *Do peer group members outperform individual borrowers?: A test of peer group lending using Canadian micro-credit data* (pp. 2003-33). Wellington: Bank of Canada. Retrieved from <https://www.bankofcanada.ca/wp-content/uploads/2010/02/wp03-33.pdf>
- Heckman, J. J. (1977). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46(4), 931-959. <https://doi.org/10.2307/1909757>
- Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, 65(2), 261-294. <https://doi.org/10.1111/1467-937x.00044>
- Herzenstein, M., Dholakia, U. M., & Andrews, R. L. (2011). Strategic herding behavior in peer-to-peer loan auctions. *Journal of Interactive Marketing*, 25(1), 27-36. <https://doi.org/10.1016/j.intmar.2010.07.001>
- Hulme, M. K., & Wright, C. (2006). Internet based social lending: Past, present and future. *Social Futures Observatory*, 11, 1-115. Retrieved from <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. (2009). *Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending?* (NBER Working Paper No. 15242, NBER, Cambridge, MA). <https://doi.org/10.2139/ssrn.1570115>
- Liao, L., Ji, L., & Zhang, W. Q. (2015). Education and credit: Evidence from P2P lending platform. *Journal of Finance Research*, 3, 146-159. Retrieved from http://en.cnki.com.cn/Article_en/CJFDTotal-JRYJ201503010.htm
- Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, 59(1), 17-35. <https://doi.org/10.1287/mnsc.1120.1560>
- Lin, X., Li, X., & Zheng, Z. (2017). Evaluating borrower's default risk in peer-to-peer lending: evidence from a lending platform in China. *Applied Economics*, 49(35), 3538-3545. <https://doi.org/10.1080/00036846.2016.1262526>
- Polena, M., & Regner, T. (2018). Determinants of borrowers' default in P2P lending under consideration of the loan risk class. *Games*, 9(4), 2-17. <https://doi.org/10.3390/g9040082>
- Pope, D. G., & Sydnor, J. R. (2011). What's in a picture? Evidence of Discrimination from Prosper. com. *Journal of Human resources*, 46(1), 53-92. <https://doi.org/10.1353/jhr.2011.0025>
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33-38. <https://doi.org/10.1080/00031305.1985.10479383>
- Zhang, K., & Chen, X. (2017). Herding in a P2P lending market: Rational inference OR irrational trust?. *Electronic Commerce Research and Applications*, 23, 45-53. <https://doi.org/10.1016/j.elerap.2017.04.001>

APPENDIX

Figure 1. Transactions on RenRenDai

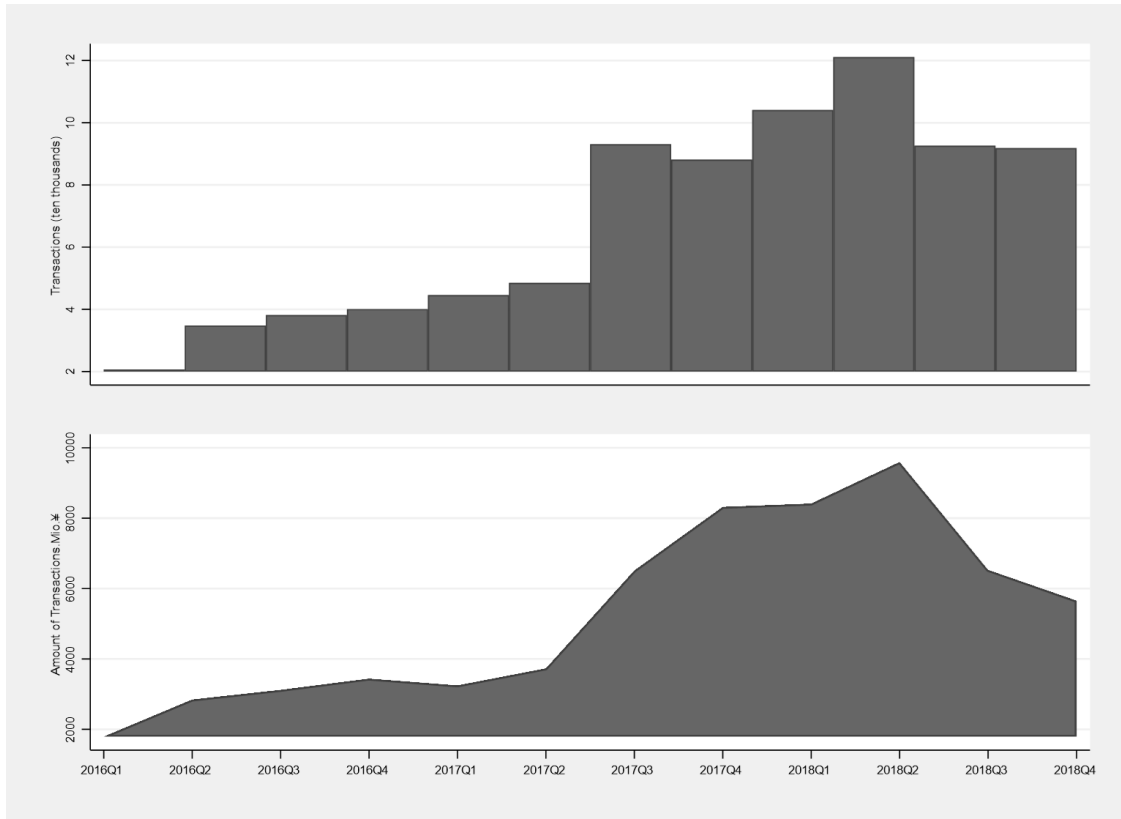


Figure 2. Distribution of borrowers by loan purpose

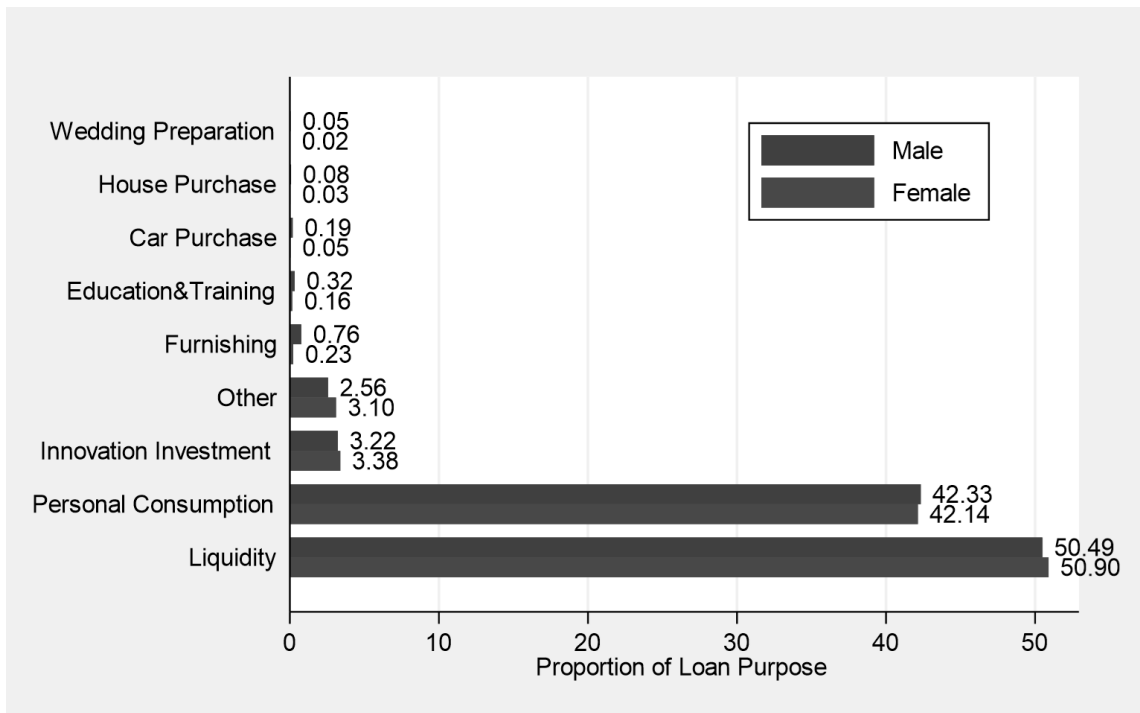


Figure 3. Data balance after propensity score match of male and female borrowers

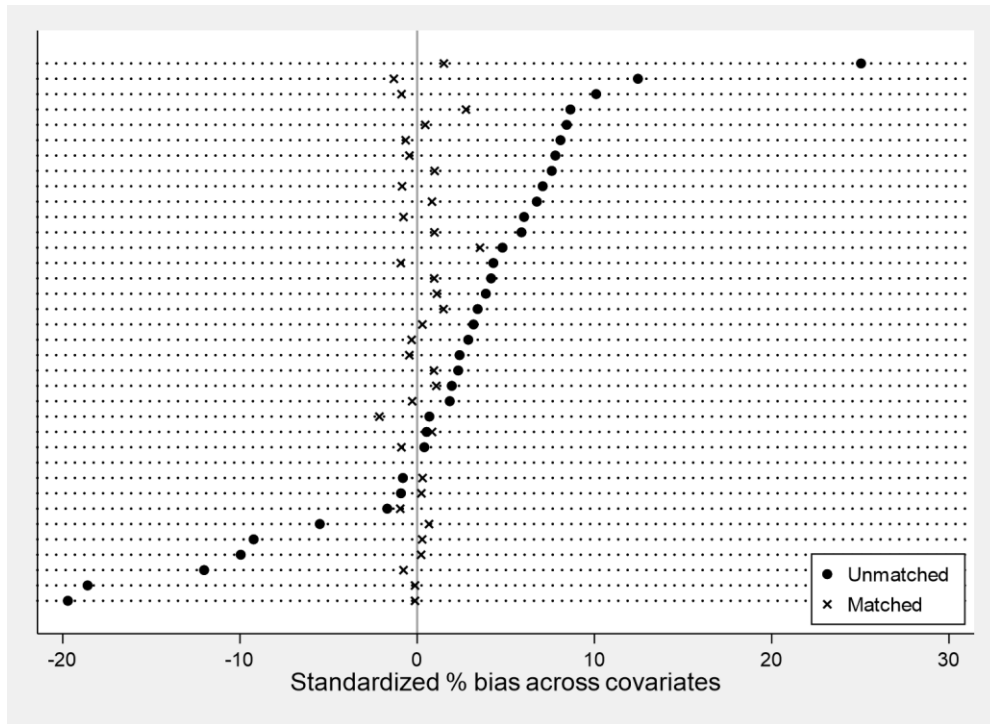


Figure 4. Common support of male and female borrowers' propensity score

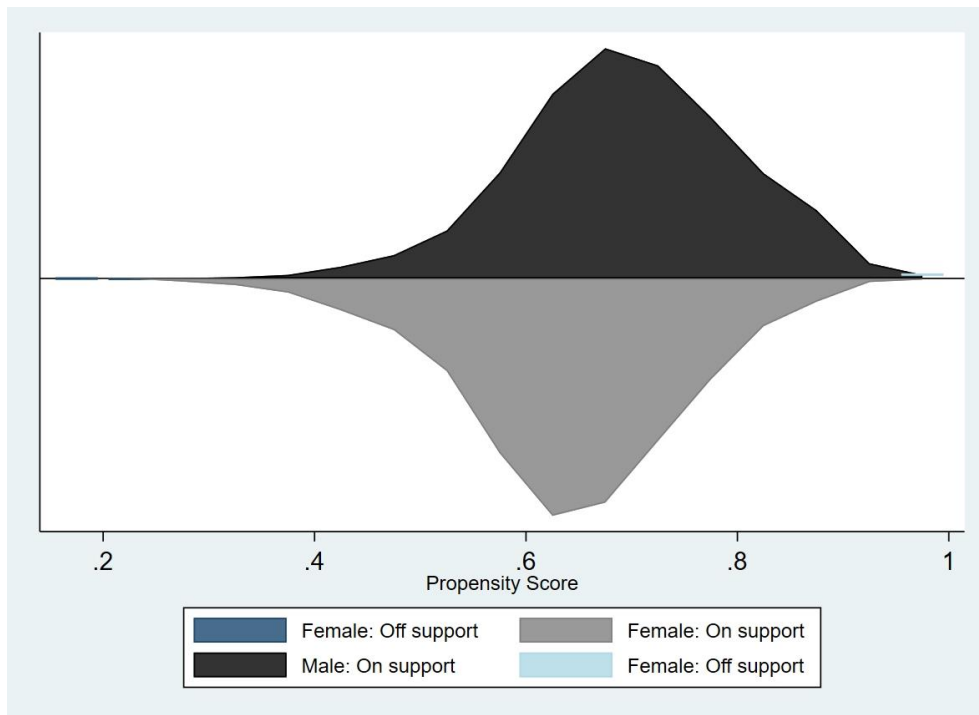


Table 1. Education level of borrowers

| <i>Education level</i> | <i>Range</i> |
|------------------------|-------------------------|
| Low | High school and below |
| Moderate | College and below |
| High | Undergraduate and below |
| Outstanding | Graduate and below |

Table 2. Ratings of borrowers and rates of risk reserve

| <i>Rating</i> | <i>AA</i> | <i>A</i> | <i>B</i> | <i>C</i> | <i>D</i> | <i>E</i> | <i>HR</i> |
|---------------|-----------|----------|----------|----------|----------|----------|-----------|
| Credit Scores | 160+ | 145-159 | 130-144 | 120-129 | 110-119 | 100-109 | 0-99 |
| Rate | 0% | 1% | 1.5% | 2% | 2.5% | 3% | 5% |

Table 3. Distribution of borrowers by gender

| | <i>Frequency</i> | <i>Percentage</i> |
|---------|------------------|-------------------|
| Females | 26,091 | 32.12 |
| Males | 55,132 | 67.88 |
| Total | 81,223 | 100.00 |

Table 4. Distribution of borrowers by monthly salary

| | <i>Monthly Salary Interval(¥)</i> | <i>Freq.</i> | <i>Percent</i> |
|---|-----------------------------------|--------------|----------------|
| 1 | ≤1,000 | 634 | 0.78 |
| 2 | ≥1,001 but ≤2,000 | 240 | 0.30 |
| 3 | ≥2,001 but ≤5,000 | 20,291 | 24.98 |
| 4 | ≥5,001 but ≤10,000 | 29,724 | 36.60 |
| 5 | ≥10,001 but ≤20,000 | 15,617 | 19.23 |
| 6 | ≥20,001 but ≤50,000 | 9,754 | 12.01 |
| 7 | ≥50,001 | 4,963 | 6.11 |
| | Total | 81,223 | 100.00 |

Table 5. Variables and definitions

| <i>Variables</i> | <i>Definition</i> |
|-------------------|--|
| Male | Male =1 if the borrower is male, otherwise Male=0. |
| Interest rate | Nominal interest rate the borrower should pay, ranging from 7-24, in % p.a. |
| Duration | Loan term, multipliers of 3 ranging from 3-36, in months. |
| Loan amount | Loan amount, divided by 1000, in RMB. |
| Education | Categorical variable with 4 values representing the borrower's education level as defined in Table X. |
| Rating | Categorical variable with 6 values based on the credit scores <i>RenRenDai</i> assigned for borrowers through risk evaluation.(See Table X) ¹ |
| Monthly salary | Categorical variable with 7 values corresponding to different monthly salary intervals described by Table X. |
| Employment status | Categorical variable representing borrowers employed by: (1) Government ; (2) Government Affiliated Institution; (3) Central State-owned Enterprise; (4) Local State-owned Enterprise; (5) World 500 Enterprise; (6)Foreign Enterprise;(7) Listed Company;(8) Private Enterprise, respectively and borrowers who are:(9) Sole Proprietor;(10) Other. |
| Age | Age of borrowers due to the loan initiation day. |
| Description | Natural logarithm of the number of characters in the borrowers' description of the loan purpose. |
| Marital status | Categorical variable corresponding to the 5 marital statuses of borrowers: Divorced, Married, Unmarried, Widowed and N/A. |
| Home ownership | If borrowers own a house, Home Ownership=1; otherwise Home Ownership=0. |
| Car ownership | If borrowers own a car, Car Ownership=1; otherwise Car Ownership=0. |
| Loan purpose | Categorical variable with 9 values describing loan purpose. |

¹ Since the possibility of default for borrowers with AA ratings is 0, our concern is restricted in the borrowers with ratings from A to HR.

Table 6. Descriptive statistics

| | <i>Mean</i> | <i>St.Dev</i> | <i>Se(Mean)</i> | <i>Variance</i> | <i>CV</i> |
|------------------------------------|-------------|---------------|-----------------|-----------------|-----------|
| Default | 0.004 | 0.059 | 0.000 | 0.003 | 16.972 |
| Male | 0.679 | 0.467 | 0.002 | 0.218 | 0.688 |
| Interest rate | 10.128 | 0.452 | 0.002 | 0.204 | 0.045 |
| Duration | 34.259 | 5.297 | 0.019 | 28.060 | 0.155 |
| Loan amount | 92.026 | 59.141 | 0.208 | 3,497.705 | 0.643 |
| <i>Education Level</i> | | | | | |
| High school | 0.115 | 0.319 | 0.001 | 0.102 | 2.777 |
| College | 0.494 | 0.500 | 0.002 | 0.250 | 1.012 |
| Undergraduate | 0.370 | 0.483 | 0.002 | 0.233 | 1.304 |
| Graduate | 0.021 | 0.142 | 0.001 | 0.020 | 6.875 |
| <i>Rating</i> | | | | | |
| A | 0.982 | 0.132 | 0.001 | 0.018 | 0.135 |
| B | 0.000 | 0.011 | - | 0.000 | 94.994 |
| C | 0.001 | 0.024 | 0.000 | 0.001 | 41.124 |
| D | 0.005 | 0.074 | 0.000 | 0.005 | 13.519 |
| E | 0.005 | 0.068 | 0.000 | 0.005 | 14.605 |
| HR | 0.007 | 0.084 | 0.000 | 0.007 | 11.874 |
| Monthly salary | 4.337 | 1.201 | 0.004 | 1.441 | 0.277 |
| <i>Employment Status</i> | | | | | |
| Local State-owned enterprises | 0.001 | 0.034 | 0.000 | 0.001 | 29.859 |
| Sole proprietor | 0.000 | 0.016 | 0.000 | 0.000 | 63.720 |
| Civil servant | 0.095 | 0.293 | 0.001 | 0.086 | 3.095 |
| Fortune global 500 companies | 0.001 | 0.030 | 0.000 | 0.001 | 33.342 |
| Government affiliated institutions | 0.149 | 0.356 | 0.001 | 0.127 | 2.388 |
| Foreign enterprises | 0.080 | 0.271 | 0.001 | 0.074 | 3.389 |
| Central state-owned enterprises | 0.189 | 0.392 | 0.001 | 0.153 | 2.072 |
| Private enterprises | 0.458 | 0.498 | 0.002 | 0.248 | 1.088 |
| Listed companies | 0.001 | 0.032 | 0.000 | 0.001 | 31.267 |
| Other | 0.026 | 0.159 | 0.001 | 0.025 | 6.114 |
| Age | 36.297 | 8.573 | 0.030 | 73.492 | 0.236 |
| Description | 4.565 | 0.184 | 0.001 | 0.034 | 0.040 |
| <i>Marital Status</i> | | | | | |
| Divorced | 0.112 | 0.316 | 0.001 | 0.100 | 2.811 |
| Married | 0.680 | 0.466 | 0.002 | 0.218 | 0.686 |
| Unmarried | 0.201 | 0.401 | 0.001 | 0.161 | 1.995 |
| Other | 0.007 | 0.082 | 0.000 | 0.007 | 12.133 |
| House ownership | 0.519 | 0.500 | 0.002 | 0.250 | 0.962 |
| Car ownership | 0.144 | 0.351 | 0.001 | 0.123 | 2.436 |
| <i>Loan Purpose</i> | | | | | |
| Liquidity | 0.506 | 0.500 | 0.002 | 0.250 | 0.988 |
| Personal consumption | 0.423 | 0.494 | 0.002 | 0.244 | 1.169 |
| Car purchase | 0.001 | 0.038 | 0.000 | 0.001 | 26.329 |
| House purchase | 0.001 | 0.025 | 0.000 | 0.001 | 39.510 |
| Marriage preparation | 0.000 | 0.020 | 0.000 | 0.000 | 50.371 |
| Education & training | 0.003 | 0.052 | 0.000 | 0.003 | 19.189 |
| Innovation investment | 0.033 | 0.178 | 0.001 | 0.032 | 5.439 |
| Furnishing | 0.006 | 0.077 | 0.000 | 0.006 | 12.970 |
| Other | 0.027 | 0.163 | 0.001 | 0.027 | 5.967 |

Table 7. Descriptive statistics (by gender)

| | <i>Female</i> | | <i>Male</i> | | <i>t-Test</i> | <i>p-Value</i> |
|------------------------------------|---------------|-----------------|-------------|-----------------|---------------|----------------|
| | <i>Mean</i> | <i>St. Dev.</i> | <i>Mean</i> | <i>St. Dev.</i> | | |
| Default | 0.001 | 0.037 | 0.004 | 0.067 | -8.400 | 0.000 |
| Interest rate | 10.118 | 0.403 | 10.133 | 0.473 | -4.731 | 0.000 |
| Duration | 34.675 | 4.460 | 34.062 | 5.640 | 16.773 | 0.000 |
| Loan amount | 96.033 | 59.868 | 90.130 | 58.700 | 13.203 | 0.000 |
| <i>Education level</i> | | | | | | |
| High school | 0.113 | 0.317 | 0.116 | 0.320 | -1.162 | 0.245 |
| College | 0.513 | 0.500 | 0.485 | 0.500 | 7.331 | 0.000 |
| Undergraduate | 0.356 | 0.479 | 0.377 | 0.485 | -5.735 | 0.000 |
| Graduate | 0.018 | 0.133 | 0.022 | 0.147 | -3.915 | 0.000 |
| <i>Rating</i> | | | | | | |
| A | 0.993 | 0.085 | 0.977 | 0.149 | 18.854 | 0.000 |
| B | 0.000 | 0.000 | 0.000 | 0.013 | -3.000 | 0.003 |
| C | 0.000 | 0.015 | 0.001 | 0.028 | -3.536 | 0.000 |
| D | 0.002 | 0.048 | 0.007 | 0.083 | -9.887 | 0.000 |
| E | 0.002 | 0.046 | 0.006 | 0.076 | -8.568 | 0.000 |
| HR | 0.003 | 0.051 | 0.009 | 0.095 | -12.859 | 0.000 |
| Monthly salary | 4.351 | 1.243 | 4.330 | 1.180 | 2.244 | 0.025 |
| <i>Employment status</i> | | | | | | |
| Local state-owned enterprises | 0.000 | 0.012 | 0.002 | 0.040 | -7.676 | 0.000 |
| Sole proprietor | 0.000 | 0.014 | 0.000 | 0.017 | -0.726 | 0.468 |
| Civil servant | 0.048 | 0.214 | 0.117 | 0.321 | -35.880 | 0.000 |
| Fortune global 500 companies | 0.001 | 0.021 | 0.001 | 0.033 | -3.331 | 0.001 |
| Government affiliated institutions | 0.172 | 0.377 | 0.139 | 0.346 | 12.090 | 0.000 |
| Foreign enterprises | 0.069 | 0.254 | 0.085 | 0.279 | -8.155 | 0.000 |
| Central state-owned enterprises | 0.157 | 0.363 | 0.204 | 0.403 | -16.890 | 0.000 |
| Private enterprises | 0.521 | 0.500 | 0.428 | 0.495 | 24.716 | 0.000 |
| Listed companies | 0.000 | 0.016 | 0.001 | 0.037 | -5.913 | 0.000 |
| Other | 0.033 | 0.177 | 0.023 | 0.150 | 7.524 | 0.000 |
| Age | 35.718 | 8.115 | 36.571 | 8.768 | -13.620 | 0.000 |
| Description | 4.564 | 0.136 | 4.565 | 0.203 | -0.907 | 0.365 |
| <i>Marital status</i> | | | | | | |
| Divorced | 0.156 | 0.363 | 0.092 | 0.288 | 25.223 | 0.000 |
| Married | 0.655 | 0.475 | 0.692 | 0.462 | -10.305 | 0.000 |
| Unmarried | 0.178 | 0.383 | 0.212 | 0.408 | -11.361 | 0.000 |
| Other | 0.010 | 0.101 | 0.005 | 0.071 | 7.558 | 0.000 |
| House ownership | 0.495 | 0.500 | 0.531 | 0.499 | -9.424 | 0.000 |
| Car ownership | 0.125 | 0.331 | 0.153 | 0.360 | -10.945 | 0.000 |
| <i>Loan purpose</i> | | | | | | |
| Liquidity | 0.509 | 0.500 | 0.505 | 0.500 | 1.089 | 0.276 |
| Personal consumption | 0.421 | 0.494 | 0.423 | 0.494 | -0.518 | 0.605 |
| Car purchase | 0.001 | 0.021 | 0.002 | 0.044 | -6.329 | 0.000 |
| House purchase | 0.000 | 0.019 | 0.001 | 0.028 | -2.630 | 0.009 |
| Marriage preparation | 0.000 | 0.012 | 0.001 | 0.023 | -2.887 | 0.004 |
| Education & training | 0.002 | 0.041 | 0.003 | 0.057 | -4.490 | 0.000 |
| Innovation investment | 0.034 | 0.181 | 0.032 | 0.177 | 1.207 | 0.227 |
| Furnishing | 0.002 | 0.048 | 0.008 | 0.087 | -11.358 | 0.000 |
| Other | 0.031 | 0.173 | 0.026 | 0.158 | 4.246 | 0.000 |

Table 8. Two-step estimation of default rate

Panel A. First-step Regression

| | <i>Interest rate</i> | | <i>Loan amount</i> | |
|------------------------------------|----------------------|---------|--------------------|----------|
| Male | 0.002 | (0.003) | -7.986*** | (0.399) |
| Duration | 0.048*** | (0.000) | 2.765*** | (0.043) |
| <i>Education level</i> | | | | |
| High School (reference category) | | | | |
| College | -0.037*** | (0.004) | 6.126*** | (0.607) |
| Undergraduate | -0.050*** | (0.004) | 22.124*** | (0.648) |
| Graduate | -0.033*** | (0.009) | 40.816*** | (1.400) |
| <i>Rating</i> | | | | |
| A (reference category) | | | | |
| B | 2.118*** | (0.113) | 47.168*** | (17.405) |
| C | 2.572*** | (0.050) | 4.188 | (7.778) |
| D | 2.241*** | (0.022) | 1.152 | (3.319) |
| E | 2.413*** | (0.022) | -2.944 | (3.451) |
| HR | 2.812*** | (0.020) | -9.222*** | (3.080) |
| Monthly salary | 0.002* | (0.001) | 9.802*** | (0.170) |
| <i>Employment status</i> | | | | |
| Other (reference category) | | | | |
| Local State-owned Enterprises | 0.220*** | (0.038) | 30.321*** | (5.779) |
| Sole proprietor | 0.248*** | (0.076) | 23.226** | (11.777) |
| Civil servant | 0.096*** | (0.009) | 40.034*** | (1.357) |
| Fortune global 500 companies | 0.240*** | (0.042) | 35.089*** | (6.392) |
| Government affiliated institutions | 0.110*** | (0.008) | 29.867*** | (1.301) |
| Foreign enterprises | 0.098*** | (0.009) | 25.905*** | (1.354) |
| Central state-owned enterprises | 0.089*** | (0.008) | 24.992*** | (1.272) |
| Private enterprises | 0.104*** | (0.008) | 11.226*** | (1.222) |
| Listed companies | 0.272*** | (0.039) | 31.787*** | (6.045) |
| Age | 0.001*** | (0.000) | 0.955*** | (0.025) |
| Description | 0.119*** | (0.009) | -8.228*** | (1.316) |
| <i>Marital status</i> | | | | |
| Other (reference category) | | | | |
| Divorced | -0.013 | (0.018) | 2.818 | (2.717) |
| Married | -0.013 | (0.017) | 5.692** | (2.679) |
| Unmarried | -0.018 | (0.018) | 2.749 | (2.729) |
| House ownership | 0.035*** | (0.002) | 4.788*** | (0.385) |
| Car ownership | -0.018*** | (0.003) | 6.571*** | (0.529) |
| <i>Loan purpose</i> | | | | |
| Other (reference category) | | | | |
| Liquidity | -0.026*** | (0.007) | -8.873*** | (1.133) |
| Personal consumption | -0.003 | (0.007) | -8.734*** | (1.137) |
| Car purchase | 0.145*** | (0.034) | -10.902** | (5.255) |
| House purchase | 0.084* | (0.049) | -20.792*** | (7.547) |
| Marriage preparation | 0.095 | (0.061) | -4.317 | (9.430) |
| Education & training | 0.868*** | (0.030) | -16.407*** | (4.545) |
| Innovation investment | 0.004 | (0.010) | -14.932*** | (1.509) |
| Furnishing | 0.148*** | (0.021) | -18.030*** | (3.210) |
| _Cons | 7.808*** | (0.046) | -68.479*** | (7.074) |
| Adj. R ² | 0.446 | | 0.232 | |

Panel B. Second-step regression

| | <i>Coefficient</i> | <i>Standard Error</i> |
|----------------------------------|----------------------|-----------------------|
| Interest Rate | -0.537 | (0.898) |
| Loan Amount | -0.004 | (0.005) |
| Male | -0.002 | (0.143) |
| Duration | 0.082 | (0.042) |
| <i>Education level</i> | | |
| High School (reference category) | | |
| College | -0.463 ^{**} | (0.142) |
| Undergraduate | -0.533 [*] | (0.207) |
| Graduate | -0.756 | (0.439) |
| <i>Rating</i> | | |
| A (reference category) | | |
| B | 4.764 [*] | (1.982) |
| C | 4.532 | (2.381) |
| D | 4.230 [*] | (2.086) |
| E | 4.954 [*] | (2.240) |
| HR | 6.328 [*] | (2.599) |
| Monthly Salary | 0.077 | (0.068) |
| Age | 0.018 | (0.010) |
| Description | -0.012 | (0.158) |
| <i>Marital status</i> | | |
| Other (reference category) | | |
| Divorced | 1.813 | (97.595) |
| Married | 1.910 | (97.595) |
| Unmarried | 1.863 | (97.595) |
| House Ownership | 0.078 | (0.107) |
| Car Ownership | -0.202 | (0.121) |
| <i>Loan purpose</i> | | |
| Other (reference category) | | |
| Liquidity | 0.012 | (0.343) |
| Personal Consumption | 0.312 | (0.352) |
| Car Purchase | 0.343 | (0.404) |
| House Purchase | 0.001 | (0.462) |
| Marriage Preparation | 0.012 | (0.487) |
| Education & Training | 0.942 | (0.930) |
| Innovation Investment | 0.481 | (0.365) |
| Furnishing | 0.455 | (0.389) |
| _cons | -3.859 | (97.857) |
| Wald-test of exogeneity | $\chi^2 = 1.824$ | $p=0.402$ |

Table 9. Determinants of default rate (part 1)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------------------------|---------------------|---------------------|----------------------|----------------------|----------------------|----------------------|
| Male | 0.054 (0.116) | 0.053 (0.114) | 0.040 (0.114) | 0.042 (0.118) | 0.041 (0.119) | 0.013 (0.121) |
| Interest rate | 0.099 (0.218) | 0.111 (0.217) | 0.126 (0.225) | 0.116 (0.214) | 0.118 (0.217) | 0.095 (0.214) |
| Duration | 0.044 (0.033) | 0.042 (0.032) | 0.039 (0.034) | 0.040 (0.032) | 0.040 (0.033) | 0.042 (0.033) |
| Loan amount | -0.000 (0.002) | -0.001 (0.002) | -0.000 (0.002) | -0.000 (0.002) | -0.000 (0.002) | 0.000 (0.002) |
| <i>Rating</i> | | | | | | |
| A (reference category) | | | | | | |
| B | 3.386** (1.140) | 3.425** (1.158) | 3.310** (1.179) | 3.374** (1.122) | 3.379** (1.126) | 3.272** (1.118) |
| C | 2.860* (1.160) | 2.887* (1.209) | 2.924* (1.252) | 2.992* (1.203) | 2.977* (1.217) | 2.926* (1.223) |
| D | 2.677* (1.114) | 2.679* (1.158) | 2.759* (1.210) | 2.844* (1.159) | 2.828* (1.174) | 2.837* (1.175) |
| E | 3.330** (1.155) | 3.338** (1.202) | 3.398** (1.261) | 3.484** (1.205) | 3.470** (1.221) | 3.494** (1.218) |
| HR | 4.505*** (1.189) | 4.491*** (1.234) | 4.539*** (1.293) | 4.644*** (1.235) | 4.631*** (1.250) | 4.654*** (1.248) |
| Monthly salary | | 0.030 (0.054) | 0.049 (0.054) | 0.047 (0.056) | 0.047 (0.056) | 0.043 (0.057) |
| Age | | 0.017** (0.006) | 0.015* (0.006) | 0.017* (0.007) | 0.016* (0.007) | 0.016* (0.007) |
| Description | | -0.062 (0.102) | -0.036 (0.105) | -0.044 (0.104) | -0.043 (0.104) | -0.062 (0.106) |
| House ownership | | 0.005 (0.093) | 0.081 (0.093) | 0.100 (0.094) | 0.096 (0.097) | 0.049 (0.102) |
| Car ownership | | -0.216* (0.107) | -0.210* (0.106) | -0.204 (0.109) | -0.208 (0.110) | -0.215 (0.110) |
| <i>Education level</i> | | | | | | |
| High School (reference category) | | | | | | |
| College | | | -0.459*** (0.113) | -0.445*** (0.117) | -0.447*** (0.117) | -0.478*** (0.119) |
| Undergraduate | | | -0.595*** (0.135) | -0.553*** (0.144) | -0.556*** (0.144) | -0.580*** (0.144) |
| Graduate | | | -0.858* (0.423) | -0.850* (0.432) | -0.849* (0.431) | -0.927* (0.426) |
| <i>Employment status</i> | | | | | | |
| Other (reference category) | | | | | | |
| Local State-owned enterprises | | | | -0.474 (0.395) | -0.473 (0.391) | -0.434 (0.387) |
| Sole proprietor | | | | -0.932 (0.599) | -0.935 (0.598) | -0.951 (0.590) |
| Civil servant | | | | -0.688 (0.368) | -0.684 (0.364) | -0.660 (0.354) |
| Fortune global 500 companies | | | | -0.716 (0.399) | -0.718 (0.398) | -0.685 (0.390) |
| Government affiliated institutions | | | | -0.663 (0.374) | -0.663 (0.370) | -0.692 (0.365) |
| Foreign enterprises | | | | -0.378 (0.361) | -0.380 (0.360) | -0.373 (0.349) |
| Central State-owned Enterprises | | | | -0.474 (0.362) | -0.472 (0.360) | -0.449 (0.351) |
| Private enterprises | | | | -0.57 (0.345) | -0.573 (0.342) | -0.580 (0.336) |
| Listed companies | | | | -0.652 (0.384) | -0.656 (0.383) | -0.654 (0.376) |
| <i>Marital status</i> | | | | | | |
| Other (reference category) | | | | | | |
| Divorced | | | | | 1.937*** (0.260) | 1.952*** (0.475) |
| Married | | | | | 2.025*** (0.230) | 2.040*** (0.461) |
| Unmarried | | | | | 2.003*** (0.271) | 2.019*** (0.480) |

Table 10. Determinants of default rate (part 2)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| <i>Loan purpose</i> | | | | | | |
| Other (reference category) | | | | | | |
| Liquidity | | | | | | 0.064 (0.272) |
| Personal consumption | | | | | | 0.369 (0.282) |
| Car purchase | | | | | | 0.291 (0.307) |
| House purchase | | | | | | 0.022 (0.364) |
| Marriage preparation | | | | | | -0.029 (0.387) |
| Education & training | | | | | | 0.483 (0.404) |
| Innovation investment | | | | | | 0.561 (0.299) |
| Furnishing | | | | | | 0.453 (0.269) |
| Constant | -6.606*** (1.038) | -7.076*** (0.972) | -6.975*** (0.994) | -6.443*** (0.884) | -8.451*** (1.060) | -8.380*** (1.240) |
| N | 81223 | 81223 | 81223 | 81223 | 81223 | 81223 |
| pseudo R ² | 0.716 | 0.719 | 0.726 | 0.728 | 0.728 | 0.733 |