ADBI Working Paper Series

ADVERSE SELECTION AND CREDIT CERTIFICATES: EVIDENCE FROM A P2P PLATFORM

Maggie Rong Hu, Xiaoyang Li,
and Yang Shi

No. 942
April 2019

Asian Development Bank Institute

Maggie Rong Hu is assistant professor at the Chinese University of Hong Kong. Xiaoyang Li is a doctoral student at the Chinese University of Hong Kong. Yang Shi is a master's student at the University of Bath, United Kingdom.

The Working Paper series is a continuation of the formerly named Discussion Paper series; the numbering of the papers continued without interruption or change. ADBI's working papers reflect initial ideas on a topic and are posted online for discussion. Some working papers may develop into other forms of publication.

Suggested citation:

Hu, M. R., X. Li, and Y. Shi. 2019. Adverse Selection and Credit Certificates: Evidence from a P2P Platform. ADBI Working Paper 942. Tokyo: Asian Development Bank Institute. Available: https://www.adb.org/publications/adverse-selection-credit-certificates-evidence-p2p-platform

Please contact the authors for information about this paper.

Email: maggiehu@cuhk.edu.hk, xiaoyangli@link.cuhk.edu.hk, yang.shi@bath.edu

Asian Development Bank Institute
Kasumigaseki Building, 8th Floor
3-2-5 Kasumigaseki, Chiyoda-ku
Tokyo 100-6008, Japan

Tel: +81-3-3593-5500
Fax: +81-3-3593-5571
URL: www.adbi.org
E-mail: info@adbi.org

Abstract

Certificates are widely used as a signaling mechanism to mitigate adverse selection when information is asymmetric. To reduce information asymmetry between lenders and borrowers, Chinese peer-to-peer (P2P) lending platforms encourage borrowers to obtain various kinds of credit certificates. As P2P markets continue to develop, it is plausible that certification may play a pivotal role in ensuring investment efficiency. We perform the first empirical investigation of this issue, using unique data from Renrendai, one of the People's Republic of China's largest P2P lending platforms. We find that surprisingly, loans with more credit certificates experience a higher rate of delinquency and default. However, lenders remain attracted by higher certificates despite lower loan performance ex post, which results in distorted capital allocation and reduced investment inefficiency. Overall, we document a setting where credit certificates fail to serve as an accurate signal due to their costless nature, where poor-quality borrowers use more certificates to boost their credit profiles and improve their funding success. Possible explanations for this phenomenon include differences in marginal benefit of certificates for different borrower types, bounded rationality, cognitive simplification, and borrower myopia.

Keywords: P2P lending, credit allocation, adverse selection, certificate, bounded rationality, cognitive simplification

JEL Classification: G10, G20, G21, G23, G40

Contents

1. INTRODUCTION

Signaling can alleviate information asymmetry and mitigate adverse selection (Akerlof, 1970; Spence, 1973; Riley, 1979; Crawford and Sobel, 1982; Austen-Smith and Banks, 2000; among others). Certificates have been widely used to signal quality when information is asymmetric. For instance, job seekers signal their professional capacity by obtaining more educational certificates (Spence, 1973). Sustainable goods producers signal their uniqueness by acquiring third-party sustainability certificates (Brach et al., 2018). Gaining access to bank loans from financial institutions is another kind of certificate that signals the promising prospects of a firm (James, 1987; Lummer and McConnell, 1989; Best and Zhang, 1993). The corporate finance literature also documents that syndicated loans (Focarelli et al., 2008; Godlewski and Sanditov 2018), project loans (Gatti et al., 2013), stapled loans (Aslan and Kumar, 2017) and loan renegotiation (Godlewski, 2015) have similar certificate effects. One crucial assumption in these studies is that the cost of obtaining certificates is inversely related to individuals' quality. Only under this condition can certificates serve as an effective signal to mitigate information asymmetry.

The Peer-to-Peer (P2P) lending platforms in the People's Republic of China (PRC) provide an ideal laboratory for investigating whether credit certificates help resolve the information asymmetry problem in the lending process. First, credit certificates are an important feature of Chinese online P2P lending platforms. To promote information transparency and facilitate investment decision-making, P2P platforms in the PRC encourage borrowers to obtain various kinds of credit certificates, such as borrower education level, marital status, address proof, employer, income level, car and property ownership, mortgage status, etc. Borrowers can voluntarily upload relevant documents for the platform to check by following simple instructions online. After passing the verification process, certificates are issued to the borrowers and listed on their profile pages.

Second, information asymmetry is of primary concern on online lending platforms. Lenders' ability to judge financial risk and information is crucial to the viability of these markets (Iyer et al., 2009). However, most lenders on these platforms are small retail investors who are inexperienced and relatively new to the investment products. They have little knowledge about either the borrower quality or the quality of the loan. In this case, they may refer to these certificates as a convenient guide for quality, based on their past experiences (Gick and Holyoak, 1980; Duhaime and Schwenk, 1985; etc.).

Third, despite its rapid growth of investment and credit demand of Chinese individuals in recent years, P2P lending in the PRC is still an emerging business area, especially in terms of regulatory monitoring. The P2P industry is considered one of the riskier and less regulated parts of the PRC's $10 trillion shadow-banking system, and it has experienced a large-scale collapsed in the past.[1] Hence, the issue of asymmetric information on Chinese P2P platforms is expected to be more severe compared to more developed P2P lending platforms internationally. To date, there is scant direct evidence on whether these peer-to-peer markets can effectively screen borrowers and allocate credit. As P2P lending and crowdfunding continues to develop, it is plausible that credit certification may play a pivotal role in ensuring efficiency.

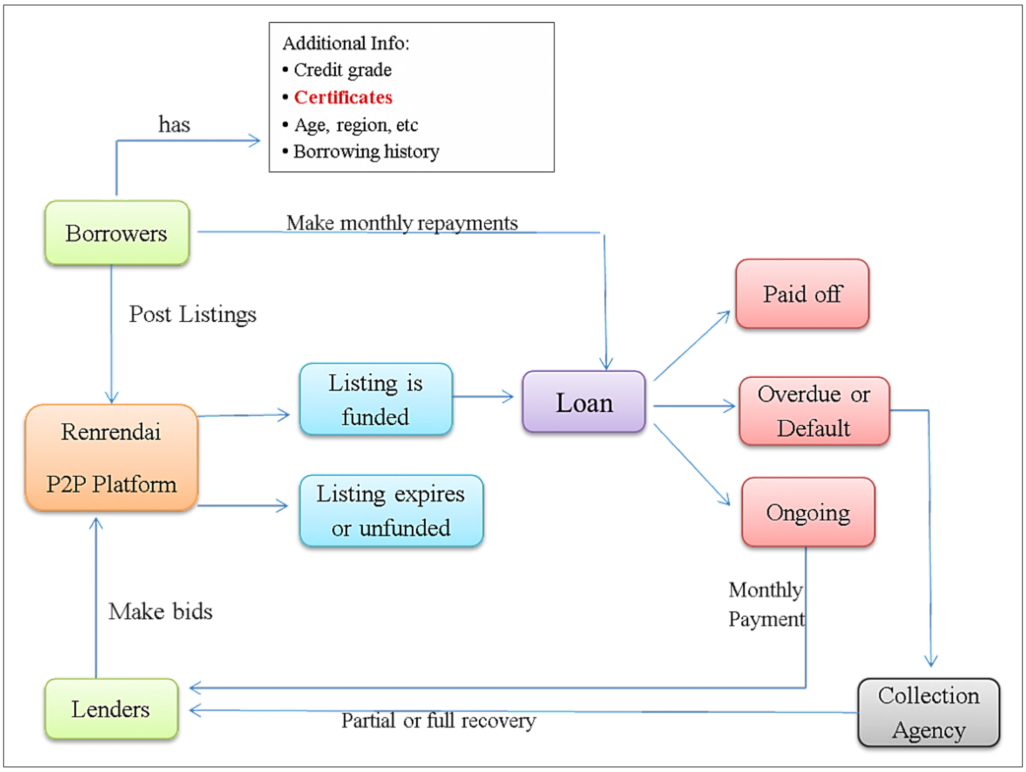
[1] Retrieved from Bloomberg "From China's Peer-to-Peer Lenders Are Falling Like Dominoes as Panic Spreads", July 20, 2018, https://www.bloomberg.com/news/articles/2018-07-20/china-s-p2p-platform-failures-surge-as-panic-spreads-in-market. In August 2016, 230 platforms collapsed. More recently, 118 Chinese P2P platforms collapsed in July 2018 alone.

This study presents the first empirical investigation of this issue, using unique data from one of the largest P2P platforms in the PRC, Renrendai (RRD). The basic lending and borrowing procedures are illustrated in Figure 1 Panel A, whereas Panel B presents the list of certificates and their frequencies in the sample. As shown in the frequency chart, platform training and ID info rank as the top two certificate types, as they are compulsory information required in order to proceed with the loan application stage. The former indicates that a user has at least skimmed through the basic policies of the platform, and the latter reveals the identity of the applicant. RRD relies heavily on mobile phone information to facilitate debt collection. In the case of delayed payments, borrowers will be contacted via phone, and thus it is not surprising to see that mobile phone is ranked as the third most frequent certificate type. The platform also has other certificates covering borrowers' income and savings, social media, contact information, third-party endorsement (e.g. onsite authentication), and formal loan application (e.g. loan description).

Using a comprehensive dataset on loan application[2] and performance from October 2010 to January 2016, we examine whether credit certificates serve as a good investment guide for capital allocation, i.e., whether borrowers with more certificates indeed have better credit quality by examining detailed loan repayment records.

Figure 1: P2P Platform Procedures and Credit Certificates

Panel A: Lending and Borrowing Process

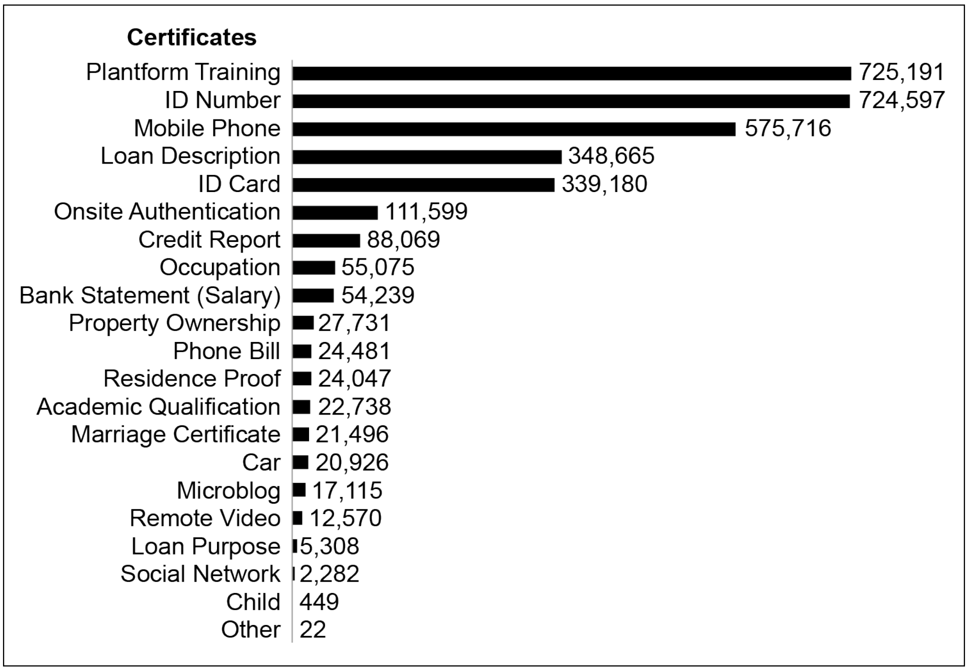


continued on next page

[2] A loan application is also known as a listing. We use these two names interchangeably.

Figure 1 *continued*

Panel B: List of Certificates



Note: Panel A presents the lending and borrowing process flow chart on Renrendai, and Panel B presents the frequencies of the certificates in our data, collected from Renrendai. The definitions of all variables are presented in Appendix 1.

To our surprise, we find that loans with more certificates actually perform worse and have a higher delinquency rate and higher default rate. Other things being equal, one additional certificate increases delinquency hazard by 16.7%. Figure 2 Panel A illustrates the hazard rate (i.e. on-time repayment without delinquency) across groups of loans with different certificate levels using the Cox proportional hazard model, which offers direct evidence that more certificates are associated with a higher hazard rate, especially at the early stage of the loan repayment.

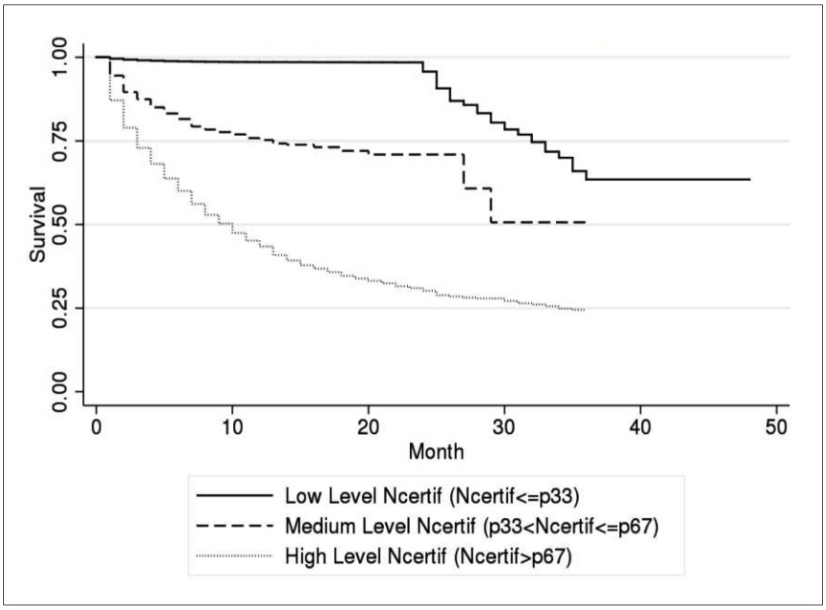
To reconcile the puzzling relationship between more certificates and higher delinquency, we next examine whether certain poor-quality borrowers self-select to obtain more certificates, to mimic the good-quality ones, to compete for funding in an opaque information environment. We analyze the determinant of certificate usage and find that indeed the number of certificates is negatively associated with the credit quality of borrowers. On average, a one-notch reduction in credit grade results in 0.337 more certificates, 3%, other things being equal.

We then check whether investors are aware of this inverse relationship between certificates and loan performance. We find that listings posted by borrowers with a higher number of credit certificates attract more capital investment. This suggests that investors are unaware of the lower performance associated with more certified loans, and therefore allocate funds to them.

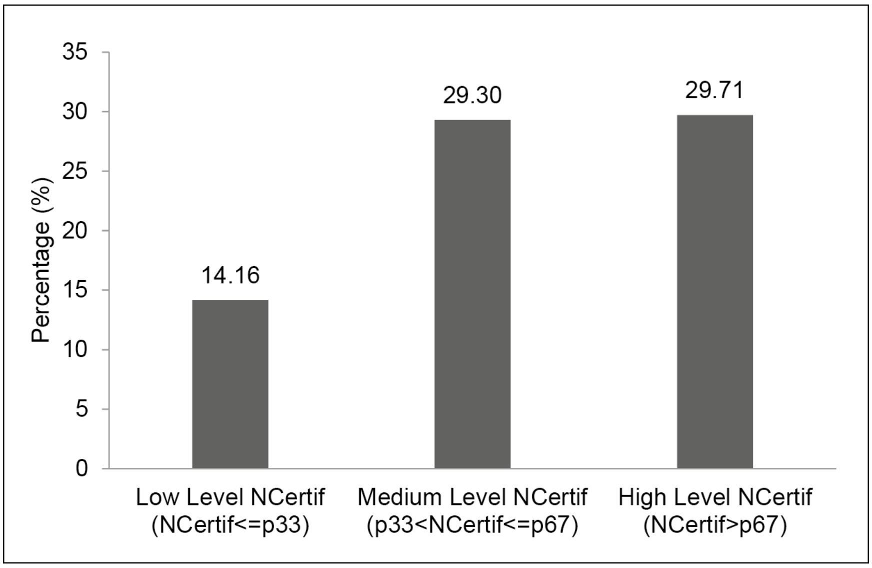
Specifically, we divide the entire sample of listings into three equal groups based on the number of certificates and compare the average funding probability of each group (Figure 2 Panel B). The funding success rate improves from 14.16% to 29.71% when a borrower moves from a low-certificate group to a high-certificate group. On average, one additional certificate increases funding success by 44.3%. controlling for loan and borrower characteristics.

Figure 2: Loan Performance and Funding Success by Certificate Level

Panel A: Loan Performance by Certificate Level



Panel B: Funding Success by Certificate Level



Note: Panel A presents the performance of loans by certificate level, and Panel B presents the funding probability by certificate level. The full sample is equally partitioned into three groups by the number of certificates obtained by borrowers. Kaplan-Meier estimators of survival function and funding probability for the high-certificate group (above 67 percentile), medium-certificate group (33 percentile and 67 percentile), and low-certificate group (below 33 percentile) are presented.

This result suggests that lenders do interpret credit certificates as a positive signal, as they are more willing to invest in listings with more certificates, even with worse performance ex post. A crucial implication for investment efficiency is that credit certificates on P2P platforms fail to serve as an effective signal in capital allocation. Capital on this platform may be allocated in an inefficient manner, whereby low-quality borrowers receive preferential treatment over high-quality borrowers.

Ideally, certificates should serve as a distinguishing mechanism so that high-quality borrowers can assert priority in obtaining funds and therefore receive preferential treatment from lenders. In our case, certificates are unable to serve their proper signaling role, as they fail to distinguish the good from the bad, resulting in losses of both lenders and high-quality borrowers. This leads to platform inefficiency; specifically, lenders take more uninformed risks without being compensated, and high-quality borrowers receive lower funding investment than they deserve. In a nutshell, credit certificates not only to fail to serve as an accurate signal in the RRD platform, but they also significantly distort credit allocation.

To ensure the robustness of our results, we conduct several additional tests. First, given that lenders tend to exercise less screening when return is guaranteed, we exclude loans with guaranteed repayment, in order to make sure our results are not affected by this type of loan. Second, to explore the linear combination of different types of certificates, we conduct principal component analysis by replacing the total number of certificates with the first main component. Third, we confirm the result on delinquency using different delinquency length in month, specifically we change the definition of delinquency and adopt different estimation techniques including single-failure model and multiple-failure model.[3] Our results remain robust after these tests, confirming the main findings that more certificates attract more investment but have worse loan performance.

Next, we provide several possible explanations for our findings on both lenders' and borrowers' behaviors. To understand why lenders rely on certificates without conducting a thorough analysis of borrowers' credit quality, we first examine their investment experiences on the P2P platform. We find that most lenders are retail investors with limited experience of P2P investment. A median lender invests in 16 loans with a median amount of RMB 450 (about USD $65) and has 6 months of investment experience on the platform.

Faced with a myriad of information about loan applications and borrower characteristics, small retail P2P lenders may find it challenging to make optimal investment decisions. Instead, they resort to cognitive simplification (Schwenk, 1984). As shown in Schunn and Dunbar (1996), earlier life experience and prior knowledge can be applied to new situations using analogical reasoning. As credit certificates are conventionally associated with better quality and favorable attributes, lenders are thus likely to reason by analogy and regard certificates as a positive signal, and thus are more willing to invest in listings with a higher number of certificates.

For borrowers, we find that low-quality borrowers use more certificates to boost their credit profile, to mimic the behavior of high-quality borrowers. However, it is puzzling why high-quality borrowers do not apply the same strategy to obtain more certificates, as they could easily be equally as successful (if not more successful) compared to low-quality borrowers. To explain this result, we argue that the marginal return of an increased number of certificates is much higher for low-quality borrowers. Although the improvement of funding likelihood is statistically significant for all borrowers, the degree of economic significance is very different. High-quality borrowers have less incentive to upload more certificates, as their funding success rate is already high enough. On the other hand, borrowers with a High Risk (HR) credit grade experience a 109.6% increase in funding success rate for each additional certificate.

In addition to this marginal benefit explanation, we can also draw on the psychology and behavioral literature. Borrowers' behavior may be subject to bounded rationality. Instead

[3] For brevity, these robustness results using different definitions of delinquency and different estimation models are omitted from the main text, but are available in Internet Appendix 2.

of always achieving optimality, given limited knowledge and attention, they make satisficing decisions; namely an action that exceeds a preset satisfactory level (Simon, 1955; Gigerenzer, 2008). While low-quality borrowers rely on additional certificates to improve their credit grade, the credit profile for high-quality individuals is typically sufficient (i.e. satisfactory), and therefore it is low-quality borrowers who actively seek to obtain more certificates.

Our research makes several important contributions. First and foremost, we document a phenomenon that credit certificates fail to serve as an accurate signal. This ineffective signaling results in distorted capital allocation and reduced investment efficiency, which contradicts the prevailing belief that credit certificates are always there to address information asymmetry and mitigate adverse selection.

Our study offers new evidence on the role of certificates in resource allocation efficiency, adding to the literature on credit certificates. A related study on certificates by Auriol and Schilizzi (2015) questions the validity of certificates, but from a totally different angle. They focus on the seed market in developing countries, where the cost of obtaining a certificate is prohibitive. In the extreme case, the high certificate cost prevents producers from obtaining it, resulting in the collapse of the certificate market.

Our study has several distinctive differences from Auriol and Schilizzi (2015). First, their theory is framed under classical economics, with rationality as an important assumption. With more assumptions on cost and market structure, they show how certificates could be less effective or useless when the cost is too high. We relax the existing assumption of rationality, and instead place more emphasis on bounded rationality.

Second, in contrast to Auriol and Schilizzi (2015), the credit certificates in our study are almost costless. Specifically, Auriol and Schilizzi (2015) demonstrate the prohibitive nature of high certificate cost, whereas we reveal how the low cost of certificates (along with bounded rationality) results in the misuse of certificates.

Further, the economic consequences also differ. In their model, when the cost drives all producers away, certificates become useless and the market deteriorates into an initial pooling equilibrium, without any additional harm caused. In the P2P lending market, the certificate itself distorts the market by attracting more funds to low-quality borrowers, thus reducing credit allocation efficiency.

This paper also contributes to the literature on P2P lending. Existing studies mainly focus on funding success rate, interest rate, default probability, and investors' bidding behavior (Duarte et al., 2012; Zhang and Liu, 2012; Lin et al., 2013; Liu et al., 2015; Dorfleitner et al., 2016, among others). We are able to observe monthly loan repayment along with overdue information in case of delinquency, thanks to the high granularity of data records. This enables us to conduct loan repayment analysis on a monthly basis using proportional hazard models. Our study also provides evidence on bounded rationality and cognitive simplification as possible causes of agents' behaviors on the P2P lending platform. Although our paper focuses on P2P lending, our conclusions also apply to other markets with amateur participants, who adopt cognitive simplification and accept certificates as a sign of good quality without question.

This paper offers practical implications for platform operator and policymakers in the P2P industry. For the past several years, P2P lending in the PRC has suffered massive defaults, with many failed attempts by various platforms to alleviate information asymmetry. Our paper reveals that, among all other possible causes, bounded rationality can be a major driving force and a unique feature of retail investors. As such, P2P platforms should consider investor attributes, such as bounded rationality, when designing their certificate processes.

This paper proceeds as follows. Section 2 introduces the institutional background and Section 3 describes the data and sample. Section 4 presents the empirical results, revealing the influence of certificates on loan performance. Section 5 investigates the relationship between a borrower's credit quality and number of certificates he or she obtains, thereby presenting the prevalent adverse selection in credit certificates. Section 6 focuses on lenders' behavior and examines how certificates affect funding success. Robustness tests are presented in Section 7. Section 8 discusses the possible channels of our findings and Section 9 concludes.

2. INSTITUTIONAL BACKGROUND

Established in 2010, RRD is one of the earliest and largest P2P lending platforms in the PRC. To raise funds from this platform, a borrower needs to create an account by offering a telephone number that can receive an SMS verification code. Basic information, such as a personal identification number, is also required. After receiving borrowers' information, the platform follows up with authenticity checks of materials provided and comprehensive assessment on borrowers' credit quality. Borrowers who pass this verification process are eligible for post-loan requirements (aka loan listings), whereas those who provide fake or suspicious materials will be denied by the platform. The detailed lending and borrowing process is illustrated in Figure 1 Panel A.

All borrowers are classified into seven different credit grades by the platform, namely: AA, A, B, C, D, E, and HR, based on the Renrendai's proprietary credit rating system. To improve their credit profile, borrowers can voluntarily provide more information to obtain certificates from the platform. RRD offers a wide array of certificates, including education level, marital status, income level, car and property ownership, housing loan, car loan, job industry, company size, job length, job type, job province, credit history, social media, and address, etc. To obtain each type of certificate, the applicant needs to prepare a list of required documents, according to platform guideline.

Credit certificates serve as an effective channel for the platform to collect additional information about an applicant. The total number of certificates is defined as NCertif. Certificates that reveal an applicant's personal information, financial status, and credit history help the platform in risk management and debt collection in case of default. They are thus defined as important certificates.[4] The total number of the important certificates is defined as NCertif_Impt.

The amount of voluntarily disclosed information shows the eagerness of a borrower to improve his or her credit profile. As certificates on platform training, ID number, mobile phone, loan description and ID card are compulsorily required, we define the rest of the certificates as voluntarily applied. The total number of the important certificates is defined as NCertif_Volun.

To obtain a certificate, an applicant simply needs to upload the required materials, most of which are as simple as photocopies of documents. Taking car certification as an example, the applicant is asked to provide: 1) a photo of vehicle license, and 2) a photo of the applicant standing beside their car, with plate number being clear and identifiable. The platform staff will then manually check the materials and decide if the applicant should receive the certificate. With the help of technology, an applicant can obtain this

[4] Specifically, important certificates include ID number, mobile phone, ID card, onsite authentication, credit report, occupation, bank statement, property ownership, car, phone bill, residence proof, and remote video.

certificate without much hassle, as most of the procedures are executed online. Therefore, the certificate process is almost costless.

Once a loan application is posted, it becomes an ongoing listing, whose real-time funding percentage, bidding time remaining, loan characteristics, and borrower characteristics are available to all lenders for their reference. The bidding is on a first-come, first-served basis, where lenders can bid on any ongoing listing that is not fully funded. Once the total amount funded reaches the initial requested amount, the bidding process closes immediately.

After getting funded, borrowers need to repay the loan on a monthly basis. There are two main repayment methods, namely, fixed monthly repayment and straight-term repayment. While the interest and principal are equally amortized for the loan duration across each month for the former, the latter requires a monthly payment of interest plus a one-time repayment of principal at the end of the borrowing period.

3. DATA DESCRIPTION

3.1 Summary Statistics

The sample includes 742,292 loan listings on RRD from October 2010 to January 2016, with 163,152 funded and 579,140 unfunded. Both borrower characteristics and loan characteristics are available for all listings. With detailed monthly repayment data for each funded loan, we are able to accurately identify all types of payment issues (i.e. overdue payments and defaults) and pinpoint the exact time when they occur. Also, the repayment data allows for multiple delinquencies within one loan. For example, a borrower may accrue several overdue payments over the course of their loan.

Table 1 provides summary statistics of variables used in our analysis. All of the data is collected from RRD, and detailed definitions of variables are presented in Appendix 1. Our focal variable, NCertif, has large variation among borrowers. On average, borrowers on the platform obtain 4.313 certificates. A similar pattern is observed for NCertif_Impt and NCertif_Volun, which vary from 0 and 11 with a mean of 2.773 and 0.658 respectively. While some borrowers obtain no certificates at all, the maximum number of obtainable certificates is 16. Forty percent of the borrowers own assets such as cars or houses, and the credit grade for the median borrower is the lowest, HR. The median borrower is 31 years old, post-tertiary educated, earns RMB 5,000 to RMB 10,000 a month from their employment, and has 1 to 3 years of working experience.

Next we look at loan characteristics. A median loan has 18 months in duration. While the maximum loan amount is as high as RMB 3 million, the smallest is only RMB 1,000, and the median amount is around RMB 40,000. Getting financed via RRD can be costly, as the average interest rate is 13.11%, and the average interest premium is 7.38%. Of the loans, 96.9% are unsecured; 15.8% have additional onsite authentication by the platform. The rest (3.1%) are secured either by guarantee or collateral.[5]

Delinquency is not common on the platform, as 96.3% of the funded loans are completely repaid on time. Among the 3.7% delinquent loans, 1.2% are repaid, but with overdue records, and the rest (2.5%) are unrepaid (i.e. default).

[5] The detailed guarantee and collateral information is presented in Table 7.

Table 1: Summary Statistics

Panel A: Borrower Characteristics (Full Sample)

Variable	N	mean	sd	p25	p50	p75	min	max
NCertif	742,292	4.313	1.642	3	4	5	0	16
NCertif_Impt	742,292	2.773	1.318	2	2	3	0	11
NCertif_Volun	742,292	0.658	1.264	0	0	1	0	11
CreditGrade	742,292	1.988	1.957	1	1	1	1	7
Age	742,276	33.529	7.373	28	31	37	18	89
EduLevel	670,294	1.857	0.780	1	2	2	1	4
JobIncomeLevel	594,206	4.068	1.218	3	4	5	1	7
JobLength	560,552	2.168	1.039	1	2	3	1	4
Single_dummy	723,459	0.521	0.500	0	1	1	0	1
Top20Province	560,663	0.562	0.496	0	1	1	0	1
HasAsset	742,292	0.400	0.490	0	0	1	0	1
HasLoan	742,292	0.166	0.372	0	0	0	0	1
NPriorLoan_Applied	742,291	2.416	3.741	1	1	3	1	148

Panel B: Loan Characteristics (Full Sample)

Variable	N	mean	sd	p25	p50	p75	min	max
Loan Amount (k)	742,292	59.648	86.885	12.000	40.000	62.000	1.000	3,000
Interest Premium	742,039	7.376	2.547	6.000	7.000	7.750	–3.100	19.540
Loan Duration (month)	742,292	17.689	10.005	12.000	18.000	24.000	1.000	48.000
Interest Rate	742,292	13.113	2.674	12.000	13.000	13.200	3.000	24.400
Secured	742,292	0.969	0.172	1	1	1	0	1
Unsecured	742,292	0.031	0.172	0	0	0	0	1

Panel C: Loan Performance (Subsample of Funded Loans)

Variable	N	mean	sd	p25	p50	p75	min	max
BadDebt	163,152	0.062	0.329	0	0	0	0	2
BadDebt (=0)	163,152	0.963	0.188	1	1	1	0	1
BadDebt (=1)	163,152	0.012	0.107	0	0	0	0	1
BadDebt (=2)	163,152	0.025	0.156	0	0	0	0	1

Note: Panel A reports the summary statistics of borrower characteristics. Panels B and C report the loan characteristics and repayment performance, respectively. A subsample of fully funded loans is used in Panel C. The definitions of all variables are presented in Appendix 1.

3.2 Univariate Analysis

3.2.1 Funded vs Unfunded

We next look at the differences in borrower and loan characteristics between funded and unfunded listings. In Panel A of Table 2, we report the number of observations and variable means in each group, and the mean differences are also presented with t-test statistics. The total records of funded and unfunded listings are 163,152 and 579,140 respectively.

Table 2: Univariate Test

Panel A: Funded and Unfunded Loan Listings (Full Sample)

	Funded Listings		Unfunded Listings		Diff in Mean
Variable	N	Mean	N	Mean	t-stat
Borrower Characteristics					
NCertif	163,152	5.122	579,140	4.085	1.037***
NCertif_Impt	163,152	3.934	579,140	2.446	1.489***
NCertif_Volun	163,152	1.905	579,140	0.306	1.599***
CreditGrade	163,152	5.360	579,140	1.038	4.322***
Age	163,149	38.417	579,127	32.152	6.265***
EduLevel	163,144	1.987	507,150	1.815	0.172***
JobIncomeLevel	163,145	4.504	431,061	3.903	0.601***
JobLength	162,952	1.737	397,600	2.344	–0.607***
Single_dummy	163,152	0.289	560,307	0.589	–0.300***
Top20Province	162,563	0.554	398,100	0.566	–0.012***
HasAsset	163,152	0.571	579,140	0.352	0.219***
HasLoan	163,152	0.320	579,140	0.122	0.198***
NPriorLoan_Applied	163,152	1.837	579,139	2.580	–0.742***
Loan Characteristics					
Loan Amount (k)	163,152	55.067	579,140	60.937	–5.869***
Interest Premium	163,074	6.357	578,965	7.663	–1.306***
Loan Duration (month)	163,152	24.005	579,140	15.909	8.096***

Panel B: High Level and Low Level Certificate Loan Listings (Full Sample)

	High NCertif		Low NCertif		Diff in Mean
Variable	N	Mean	N	Mean	t-stat
Borrower Characteristics					
Funding Success	262,647	0.295	479,645	0.179	0.116***
CreditGrade	262,647	2.203	479,645	1.870	0.333***
Age	262,647	34.129	479,629	33.201	0.928***
EduLevel	262,300	1.941	407,994	1.802	0.139***
JobIncomeLevel	261,938	4.162	332,268	3.993	0.169***
JobLength	261,568	2.231	298,984	2.112	0.119***
Single_dummy	262,610	0.455	460,849	0.570	–0.135***
Top20Province	260,731	0.551	299,932	0.572	–0.021***
HasAsset	262,647	0.564	479,645	0.310	0.254***
HasLoan	262,647	0.239	679,645	0.125	0.114***
NPriorLoan_Applied	262,647	3.915	479,644	1.596	2.320***
Loan Characteristics					
Loan Amount (k)	262,634	61.568	479,640	58.597	2.971***
Interest Premium	262,554	7.281	479,485	7.428	–0.148***
Loan Duration (month)	262,647	18.712	479,645	17.128	1.584***

continued on next page

Table 2 *continued*

Panel C: Loan Listings with High Level and Low Level Certificates (Subsample of Funded Loans)

	High NCertif		Low NCertif		Diff in Mean
Variable	**N**	**Mean**	**N**	**Mean**	**t-stat**
Borrower Characteristics					
CreditGrade	77,431	4.834	85,721	5.835	–1.001***
Age	77,431	37.322	85,718	39.405	–2.083***
EduLevel	77,423	2.015	85,721	1.961	0.054***
JobIncomeLevel	77,424	4.458	85,721	4.546	–0.089***
JobLength	77,252	1.785	85,700	1.694	0.092***
Single_dummy	77,431	0.300	85,721	0.279	0.022***
Top20Province	76,867	0.534	85,696	0.571	–0.037***
HasAsset	77,431	0.638	85,721	0.511	0.127***
HasLoan	77,431	0.349	85,721	0.294	0.055***
NPriorLoan_Applied	77,431	2.577	85,721	1.169	1.408***
Loan Characteristics					
Loan Amount (k)	77,431	52.585	85,721	57.310	–4.725***
Interest Premium	77,398	6.320	85,676	6.390	–0.071***
Loan Duration (month)	77,431	21.770	85,721	26.024	–4.254***
Loan Performance					
BadDebt	77,431	0.116	85,721	0.012	0.104***
BadDebt (=0)	77,431	0.932	85,721	0.992	–0.060***
BadDebt (=1)	77,431	0.021	85,721	0.003	0.017***
BadDebt (=2)	77,431	0.048	85,721	0.005	0.043***

Note: Panel A divides the full sample into funded and unfunded groups, demonstrating differences in borrower and loan characteristics, while Panel B equally partitions the full sample by number of certificates. The number of observations, sample mean, difference in mean, and t-test significance are presented. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

Consistent with our hypothesis, the average number of certificates for the funded listings is 5.122, which is significantly higher than that for unfunded listings. The number of important certificates and voluntarily applied certificates also differ significantly between the funded and unfunded loans. While the funded loans have 3.934 important certificates and 1.905 voluntarily applied certificates on average, the means for unfunded loans are 2.446 and 0.306. Also, borrowers of funded loans have positive attributes, such as more advanced education, higher income, possession of assets, and higher credit grades.

Theoretically, the impact of having car and house loans is twofold. On the one hand, the existing liability increases the leverage of the borrower. On the other hand, as suggested by the banking literature, access to bank loans certifies banks' trust toward the borrower, thus signaling good quality (James, 1987; Lummer and McConnell, 1989; Best and Zhang, 1993). Our data indicates that the signaling effect plays the dominant role, as the funding success rate for borrowers with bank loans is higher. For loan characteristics, funded loans have a smaller loan amount, lower interest premium, and longer duration.

3.2.2 High Certificate vs Low Certificate

Table 2 Panel B presents the differences in loans between high and low certificate levels, where the funded loans are equally partitioned into two groups by the number of certificates. In general, the high certificate level is associated with relatively better borrower attributes and more favorable loan terms.

Specifically, borrowers with high certificate levels attain more advanced education and earn a higher income. They also possess more assets compared to low-certificate borrowers and are more likely to receive loans from financial institutions. Further, their track records indicate that they have previously applied for more loans. Borrowers with a high certificate level apply for 3.915 loans on average, which is around 2.45 times higher than that of borrowers with a low certificate level. In terms of cost, borrowers with a high certificate level also experience lower financing costs: with a mean of 7.281, they pay around 15 basis points less interest premium compared to low-certificate borrowers.

We focus on the funded subsample in Panel C. Notably, borrowers with a high certificate level have, on average, a lower credit grade, opposite to the patterns in Panel B, where high levels of certificates are associated with better credit ratings. The relationship between credit rating and number of certificates is driven by two distinct forces. On one hand, the self-selection effect suggests that borrowers with a worse credit profile choose to get more certificates; on the other hand, certificates can boost credit profile, so they have an elevating effect on credit profile. The elevating effects differ among borrowers. While a low-quality borrower can boost his or her credit profile using more certificates, it is less rewarding for a high-quality one to do so.

As shown in Panel A, the average credit grade for funded borrowers is much higher than for those of the unfunded ones; hence, the marginal benefit on funding success is much weaker for the funded group. Consequently, the negative relationship in Panel C is mainly driven by the self-selection effect that we are interested in. The positive correlation in Panel B, however, captures a mixture of the aforementioned two effects. More importantly, the differences between funded sample and full sample also indicate that it is the low-quality borrowers that substantially boost their credit profile via certificates. Section 5 uses the funded subsample and analyzes this issue in depth.

The funded sample also allows us to investigate the performance of loans of different certificate level. We find that high certificate loans have significantly higher delinquencies. Specifically, compared with the low certificates group, loans with high certificates have, on average, 8.6 times more delinquencies and 9.6 times higher default rates.

4. CERTIFICATE AND LOAN PERFORMANCE

In this section, we formally analyze how certificates affect the loan performance using multivariate regression, controlling for other relevant factors, such as loan characteristics, borrower characteristics, and borrower track records.

The analysis starts from ordered logit regressions, where on-time repayment (=0), overdue repayment (=1), and default (=2) are modeled. A duration analysis is then adopted to further take the multiple delinquencies within one loan into consideration.

4.1 Ordered Logit Models

We first analyze loan repayments using an ordered logit regression, where all loans are classified into three ordered categories, namely: repaid (=0, loans repaid on time for each period), overdue (=1, loans with delayed payment records that are eventually repaid fully at the maturity of the loan), and default (=2, loans that are not fully repaid at the maturity of the loan). Table 3 reports the coefficients from the ordered logit regressions with standard errors in parentheses. The dependent variable, BadDebt, equals 2 if the loan is defaulted, 1 if the loan is repaid with overdue records and 0 if the loan is on time repaid for every period.

Table 3: Number of Certificates and Delinquency

Dependent Variable: BadDebt	(1)	(2)	(3)	(4)	(5)	(6)
NCertif	0.048***	0.075***				
	(0.009)	(0.010)				
NCertif_Impt			0.028**	0.042***		
			(0.013)	(0.014)		
NCertif_Volun					0.002	0.030**
					(0.012)	(0.013)
CreditGrade	–1.503***	–1.520***	–1.507***	–1.525***	–1.507***	–1.528***
	(0.021)	(0.023)	(0.021)	(0.023)	(0.021)	(0.023)
logLoanAmount	0.254***	0.098***	0.256***	0.096***	0.261***	0.097***
	(0.023)	(0.028)	(0.023)	(0.028)	(0.023)	(0.028)
Interest Premium	0.072***	0.072***	0.072***	0.073***	0.073***	0.074***
	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)
Loan Duration (month)	0.039***	0.045***	0.039***	0.045***	0.038***	0.045***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		0.026***		0.026***		0.026***
		(0.003)		(0.003)		(0.003)
EduLevel		–0.386***		–0.376***		–0.379***
		(0.021)		(0.021)		(0.021)
JobIncomeLevel		0.116***		0.123***		0.124***
		(0.016)		(0.016)		(0.016)
JobLength		0.038**		0.040**		0.040**
		(0.018)		(0.018)		(0.018)
Single_dummy		0.089**		0.076**		0.080**
		(0.038)		(0.038)		(0.038)
Top20Province		0.118***		0.121***		0.121***
		(0.033)		(0.033)		(0.033)
HasAsset		–0.015		–0.003		0.002
		(0.040)		(0.041)		(0.041)
HasLoan		–0.379***		–0.366***		–0.365***
		(0.043)		(0.043)		(0.043)
NPriorLoan_Applied		–0.020**		–0.016**		–0.017**
		(0.008)		(0.008)		(0.008)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	163,074	162,460	163,074	162,460	163,074	162,460
Pseudo R-squared	0.470	0.485	0.469	0.484	0.469	0.484

Note: This table presents the ordered logit regression results with dependent variable BadDebt; the number equals 2 if the loan is defaulted, 1 the loan is repaid with overdue records and 0 if the loan is on time repaid for every period. Specifications (1) and (2) focus on the number of total certificates, (3) and (4) focus on the number of important certificates, and (5) and (6) focus on the number of voluntarily applied certificates. Estimated coefficients are reported along with heteroskedasticity robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

The specification in Models 1, 3, and 5 incorporate only our focal variables "NCertif," "NCertif_Impt," and "NCertif_Volun" respectively, and more control variables are included in Models 2, 4, and 6. Throughout our regression analysis, year quarter fixed effects are included in all specifications to control for unserved time effects (Lin et al., 2013). Coefficients for number of total certificates and important certificates are positive and statistically significant across all models, and the number of voluntarily applied certificates also has a coefficient significantly larger than zero when the full set of control variables are included, suggesting that borrowers with more certificates, on average, are more likely to have worse performance. In Models 2, 4, and 6 with full set of controls, show that having one additional certificate, important certificate, and voluntarily applied certificates increases the odds of having a deteriorated payment record by 7.8% (=exp(0.075)-1), 4.3% (=exp(0.042)-1), and 3.0% (=exp(0.030)-1) respectively.

4.2 Cox Proportional Hazard Model

A single loan may have multiple delinquency events, e.g. multiple delayed payments within the loan duration. We next construct a variance-corrected multiple-failure Cox proportional hazards model (Andersen-Gill model) to incorporate this multiple-failure attribute into our analysis (Anderson and Gill, 1982). Compared with a single-failure Cox proportional hazards model, the Andersen-Gill model not only fully utilizes all delinquency information in the data, but also corrects the estimation of the covariance matrix by taking the correlation between delinquencies into consideration (He et al., 2019).

We present the estimated hazard ratios from the variance-corrected multiple-failure Cox proportional hazards model in Table 4. In Panel A, delinquency is defined as any payment overdue for 1 month or longer. Panel B adopts a stricter definition in identifying delinquencies, in which only consecutive overdue records for 4 months or longer are recognized (Duarte et al., 2012).[6]

Consistent with prior results, the hazard ratios for the number of total certificates, important certificates, and voluntarily applied certificates are larger than 1 and significant in all models. After controlling for other relevant factors, a one unit increase in total, important, and voluntarily applied certificate number raises the conditional probability of delinquency by 16.3%, 22.6%, and 19.4%. Defining delinquency as default or 4-month consecutive overdue payments as delinquency, these three ratios change to 14.1%, 20.0%, and 16.7% respectively.

Quantitatively, the hazard ratios are much higher than the odds ratios in the ordered logit regression. The difference between these two models is that the variance-corrected multiple-failure Cox proportional hazards model takes the number of delinquencies into consideration, which results in the larger hazard ratios. The higher hazard ratios indicate that certificates have a twofold impact on loan performance. They not only raise the probability of delinquency, but also increase the occurrences of delinquency events. To our best knowledge, this is also the first paper that uses a variance-corrected multiple-failure Cox model (Anderson and Gill, 1982) to examine the determinants of P2P lending delinquency, which not only fully utilizes the information of all delinquent events of each loan, but also adjusts the correlation between each event.

[6] As a robustness test, we present results using both the single failure model and changing the definition of delinquency to two-month consecutive overdue payments, following Lin et al. (2013). The results are qualitatively the same and are presented in Internet Appendix 2.

Table 4: Hazard Model Estimation

Panel A: One-month Overdue Payment Criterion

Dependent Variable: Default Hazard	(1)	(2)	(3)	(4)	(5)	(6)
NCertif	1.167***	1.163***				
	(0.005)	(0.005)				
NCertif_Impt			1.233***	1.226***		
			(0.008)	(0.008)		
NCertif_Volun					1.198***	1.194***
					(0.007)	(0.007)
CreditGrade	0.408***	0.418***	0.396***	0.407***	0.386***	0.398***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
logLoanAmount	1.196***	1.192***	1.194***	1.187***	1.202***	1.193***
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Interest Premium	1.084***	1.106***	1.071***	1.094***	1.063***	1.087***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
Loan Duration (month)	0.975***	0.973***	0.976***	0.974***	0.978***	0.975***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.004***		1.004***		1.004***
		(0.000)		(0.000)		(0.000)
EduLevel		0.997		1.001		0.997
		(0.005)		(0.005)		(0.005)
JobIncomeLevel		1.007*		1.009**		1.011***
		(0.004)		(0.004)		(0.004)
JobLength		1.061***		1.060***		1.064***
		(0.005)		(0.005)		(0.005)
Single_dummy		1.019**		1.021**		1.018*
		(0.010)		(0.010)		(0.010)
Top20Province		1.011		1.012		1.012
		(0.008)		(0.008)		(0.008)
HasAsset		0.977**		0.972**		0.976**
		(0.011)		(0.011)		(0.012)
HasLoan		0.916***		0.916***		0.910***
		(0.009)		(0.009)		(0.009)
NPriorLoan_Applied		1.007*		1.008*		1.003
		(0.004)		(0.004)		(0.005)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	398,790	398,105	398,790	398,105	398,790	398,105
Pseudo R-square	0.076	0.076	0.076	0.076	0.075	0.076

continued on next page

Table 4 *continued*

Panel B: Four-month Consecutive Overdue Payment Criterion

Dependent Variable: Default Hazard	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NCertif	1.144***	1.141***				
	(0.006)	(0.006)				
NCertif_Impt			1.205***	1.200***		
			(0.009)	(0.009)		
NCertif_Volun					1.169***	1.167***
					(0.008)	(0.008)
CreditGrade	0.397***	0.408***	0.386***	0.398***	0.379***	0.390***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
logLoanAmount	1.203***	1.193***	1.201***	1.191***	1.207***	1.195***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)
Interest Premium	1.072***	1.099***	1.060***	1.087***	1.051***	1.079***
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
Loan Duration (month)	0.990***	0.986***	0.990***	0.987***	0.992***	0.988***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.004***		1.004***		1.004***
		(0.000)		(0.000)		(0.000)
EduLevel		0.987**		0.990*		0.987**
		(0.005)		(0.005)		(0.005)
JobIncomeLevel		1.017***		1.018***		1.020***
		(0.004)		(0.004)		(0.004)
JobLength		1.058***		1.056***		1.058***
		(0.005)		(0.005)		(0.005)
Single_dummy		1.031***		1.032***		1.031***
		(0.010)		(0.010)		(0.010)
Top20Province		1.012		1.013		1.013
		(0.008)		(0.008)		(0.008)
HasAsset		0.985		0.982		0.986
		(0.012)		(0.012)		(0.012)
HasLoan		0.903***		0.903***		0.899***
		(0.009)		(0.009)		(0.009)
NPriorLoan_Applied		1.002		1.002		0.998
		(0.006)		(0.007)		(0.007)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	366,223	365,715	366,223	365,715	366,223	365,715
Pseudo R-square	0.067	0.068	0.067	0.068	0.067	0.067

Note: This table investigates the impact of certificate number on the conditional probability (i.e. hazard) of delinquency. In Panel A, delinquency is defined as default or overdue payments for 1 month or longer, while in Panel B, default or overdue payments for 4 months or longer is regarded as delinquency. Within each panel, Specifications (1) and (2) focus on the number of total certificates, (3) and (4) focus on the number of important certificates, and (5) and (6) focus on the number of voluntarily applied certificates. Hazard ratios from the variance-corrected multiple-failure Cox proportional hazards model (i.e. Anderson-Gill model) are reported, along with standard errors in parentheses clustered at loan level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

5. DETERMINANT OF CREDIT CERTIFICATES

To uncover the reason why certificates are inversely related to loan performance, we analyze the possible determinants of certificates related to borrower attributes. Table 5 presents the OLS estimation results with the number of certificates as the dependent variable, whereas credit grade, borrower characteristics, and borrowing experience are included as independent variables.

The number of total certificates, important certificates, and voluntarily applied certificates are used as the dependent variables in the first, middle, and last two columns of Table 5. Credit grade is the independent variable in Model 1, 3, and 5, and Model 2, 4, and 6 add other borrower characteristics and previous borrowing records as control variables. The coefficients on credit grade are significantly negative when number of total certificates and important certificates are used as dependent variable. Although statistically insignificant, the credit grade and number of voluntarily applied certificates are also inversely correlated. Quantitatively, a one-notch increase in credit grade is, on average, associated with a 0.325-unit reduction in the number of total certificates and 0.118 less important certificates.

We also find that single younger borrowers from large cities with a higher education level, higher income level, and shorter working experience tend to use more certificates. Borrowers who have housing or automobile assets are more likely to showcase these in certificates, potentially to serve as collateral or assurance. Previous borrowing experience is associated with more certificates, indicating that more experienced borrowers know better how to utilize certificates in boosting their credit profile.

The above estimates may be subject to reverse causality, as acquiring additional certificates also improves credit grades; we term this influence an elevating effect. Therefore, the estimated coefficients reflect the combination of two effects, i.e. self-selection effect (negative) and elevating effect (positive). Therefore, the impact of the pure self-selection effect should be even smaller than –0.235, indicating a stronger inverse correlation between borrower quality and number of certificates.

It is obvious why low-quality borrowers obtain more certificates, as they can use more certificates to improve their credit profiles and attract investors. However, it is still necessary to understand why high-quality individuals obtain much fewer certificates. One possible cause is bounded rationality and satisficing decision.

Similar to lenders, borrowers also have limited expertise in investment and finance, and their behaviors may also be subject to bounded rationality. Instead of pursing optimality, they make satisficing decisions, i.e. using a preset satisfactory level as a key decision criterion (Simon, 1955; Simon 1979; Gigerenzer, 2008). A borrower only cares about their funding probability and financing cost. As long as the funding success reaches a preset satisfactory level, they will not bother to obtain more certificates, even if the process is nearly costless. In contrast, low-quality applicants, being unsatisfied with their credit profiles, continue to acquire certificates until their funding outcomes become satisfactory. In the end, we observe an adverse selection in certificates; namely borrowers with poor credit profiles choose to obtain more certificates.

Table 5: Determinants of Number of Certificates

	Dependent Variable: NCertif		Dependent Variable: NCertif_Impt		Dependent Variable: NCertif_Volun	
	(1)	(2)	(3)	(4)	(5)	(6)
CreditGrade	–0.325***	–0.337***	–0.118***	–0.119***	–0.005	–0.003
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)
Age		–0.006***		–0.004***		–0.005***
		(0.000)		(0.000)		(0.000)
EduLevel		0.062***		0.006**		0.041***
		(0.005)		(0.003)		(0.003)
JobIncomeLevel		0.003		0.009***		0.008***
		(0.003)		(0.002)		(0.002)
JobLength		–0.110***		–0.050***		–0.048***
		(0.006)		(0.004)		(0.004)
Single_dummy		–0.077***		–0.028***		–0.060***
		(0.007)		(0.005)		(0.005)
Top20Province		0.103***		0.065***		0.066***
		(0.006)		(0.005)		(0.005)
HasAsset		0.435***		0.395***		0.403***
		(0.008)		(0.006)		(0.007)
HasLoan		0.067***		0.026***		0.049***
		(0.010)		(0.007)		(0.008)
NPriorLoan_Applied		0.076***		0.040***		0.060***
		(0.002)		(0.001)		(0.002)
Constant	9.474***	8.749***	4.524***	4.023***	3.282***	2.543***
	(0.215)	(0.213)	(0.111)	(0.099)	(0.169)	(0.165)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	163,152	162,538	163,152	162,538	163,152	162,538
Adj. R-squared	0.510	0.574	0.629	0.670	0.587	0.647

Note: This table presents the relationship between a borrower's credit grade and the number of certificates obtained. The independent variables are the number of total certificates, the number of important certificates, and the number of voluntarily applied certificates in specifications (1)-(2), (3)-(4), and (5)-(6) respectively. OLS regression coefficients are reported, along with heteroskedasticity robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

6. CERTIFICATES AND FUNDING SUCCESS RATE

Prior work on signaling typically assumes signals are costless and verifiable (Grossman 1981; Milgrom 1981). However, many follow-up studies have explored scenarios where the assumption that disclosures must be made truthfully is relaxed, since the seminal work of Crawford and Sobel (1982. It has shown that lenders in the P2P market believe in those "cheap-talks", i.e. costless, voluntarily disclosed, and unverified information, and make investment decisions in accordance with the costless signal (Michels, 2012).

We have so far noted that more certificates are associated with poor ex ante credit grade and higher ex post delinquency hazard in the market we study. A natural question is whether lenders are sophisticated enough to recognize the inverse relationship between certificates and loan performance, or do they still trust the certificates and simply interpret them as positive signals and invest in loans with more certificates. To answer this question, we examine the impact of the number of certificates on funding probability. Table 6 presents the logit regression model results, where the independent variable, Funding Success, is a dummy equal to 1 if the listing is funded and 0 otherwise. Panel A presents the full sample results and Panel B divides the loan applications by credit grade to reveal the heterogeneous impact of certificates.

In Panel A, Model 1, 3, and 5 only includes our focal variable, NCertif (NCertif_Impt or Ncertif_Volun), credit grade and loan characters. A full set of control variables consisting of loan characteristics, borrower characteristics, and prior borrowing records is introduced into Model 2, 4 and, 6. Logit regression coefficients are reported along with standard errors in parentheses. All of the coefficients of our focal variables are highly significant, above zero, indicating that having more certificates facilitates the fund raising. The economic significances are also quite remarkable. Other things being equal, one additional certificate, important certificate, and voluntarily applied certificate increases the funding odds by 44.3% (=exp(0.367)-1), 88.3% (=exp(0.633)-1), and 57.1% (=exp(0.452)-1).

Positive attributes, such as higher credit grade, advanced education level, higher income level, and longer working experience, are associated with higher funding success. In addition, interest premium as a comprehensive measure of risk is negatively related to funding probability. On average, a 1% increase in interest premium lowers funding odds by around 9.2% (=exp(-0.097)-1) to 10.0% (=exp(-0.105)-1) depending on specification. Prior loans from financial institutions also improve funding probability, consistent with the notion that bank loans can signal high borrower quality.

To uncover the reason why high-quality borrowers do not present a full set of certificates, we argue that it is possible that credit certificates boost funding success rate with varying degrees for high-quality and low-quality borrowers. Panel B presents the subsample regression outcomes, using the HR-rated loan sample only in the first three columns, and AA-rated loans only, which reflects the difference effect of the number of certificates on funding success for the highest and lowest quality borrowers.

We find that while certificates increase funding probability remarkably in the low credit grade group, the effect in the high rating group is very limited, indicating a diminishing effect of certificates in boosting borrower's credit quality. Take the number of total certificates as an example, while one more certificate increases funding odds by 44.3% in the full sample (Model 2 in Panel A), the effect is more than doubled in the HR subsample and reached 109.6% (=exp(0.740)-1) in Model 4 Panel B. The influence in the AA-rated sample, in contrast, is not significant.

Table 6: Certificates and Funding Success

Panel A: Certificates and Funding Success

Dependent Variable: Funding Success	(1)	(2)	(3)	(4)	(5)	(6)
NCertif	0.449***	0.367***				
	(0.004)	(0.004)				
NCertif_Impt			0.729***	0.633***		
			(0.005)	(0.006)		
NCertif_Volun					0.547***	0.452***
					(0.005)	(0.006)
CreditGrade	1.707***	1.637***	1.617***	1.553***	1.592***	1.528***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
logLoanAmount	–0.572***	–0.865***	–0.612***	–0.889***	–0.542***	–0.854***
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
Interest Premium	–0.109***	–0.100***	–0.114***	–0.105***	–0.103***	–0.097***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Loan Duration (month)	0.012***	0.020***	0.015***	0.022***	0.012***	0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age		0.037***		0.034***		0.037***
		(0.001)		(0.001)		(0.001)
EduLevel		0.187***		0.209***		0.191***
		(0.010)		(0.010)		(0.010)
JobIncomeLevel		0.316***		0.301***		0.333***
		(0.007)		(0.007)		(0.007)
JobLength		0.122***		0.092***		0.128***
		(0.008)		(0.008)		(0.008)
Single_dummy		–0.124***		–0.151***		–0.117***
		(0.017)		(0.018)		(0.017)
Top20Province		–0.183***		–0.182***		–0.187***
		(0.015)		(0.016)		(0.015)
HasAsset		–0.009		–0.104***		–0.026
		(0.019)		(0.019)		(0.019)
HasLoan		0.050**		0.005		0.039*
		(0.021)		(0.021)		(0.021)
NPriorLoan_Applied		–0.080***		–0.079***		–0.076***
		(0.001)		(0.001)		(0.001)
Constant	–7.851***	–9.066***	–6.950***	–8.199***	–5.659***	–7.338***
	(0.247)	(0.229)	(0.226)	(0.211)	(0.244)	(0.225)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	742,021	556,538	742,021	556,538	742,021	556,538
Pseudo R-squared	0.802	0.794	0.809	0.801	0.797	0.792

continued on next page

Table 6 *continued*

Panel B: Certificates and Funding Success by Credit Grade

Dependent Variable: Funding Success	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
		AA			**HR**	
NCertif	–0.051			0.740***		
	(0.037)			(0.006)		
NCertif_Impt		0.028			1.078***	
		(0.050)			(0.008)	
NCertif_Volun			–0.025			0.855***
			(0.041)			(0.007)
logLoanAmount	0.034	0.071	0.049	–0.913***	–0.943***	–0.901***
	(0.085)	(0.084)	(0.085)	(0.011)	(0.012)	(0.011)
Interest Premium	0.061*	0.058*	0.060*	–0.083***	–0.084***	–0.076***
	(0.034)	(0.034)	(0.034)	(0.004)	(0.004)	(0.004)
Loan Duration (month)	–0.061***	–0.065***	–0.063***	0.010***	0.011***	0.013***
	(0.012)	(0.012)	(0.012)	(0.001)	(0.001)	(0.001)
Age	0.019	0.026	0.022	0.029***	0.022***	0.030***
	(0.020)	(0.020)	(0.020)	(0.002)	(0.002)	(0.002)
EduLevel	0.218	0.250*	0.228*	0.174***	0.261***	0.192***
	(0.139)	(0.140)	(0.138)	(0.012)	(0.013)	(0.012)
JobIncomeLevel	0.058	0.044	0.055	0.363***	0.343***	0.371***
	(0.079)	(0.080)	(0.079)	(0.009)	(0.009)	(0.009)
JobLength	0.120	0.157	0.138	0.307***	0.296***	0.313***
	(0.113)	(0.111)	(0.113)	(0.011)	(0.011)	(0.011)
Single_dummy	–0.148	–0.030	–0.108	–0.081***	–0.157***	–0.072***
	(0.309)	(0.308)	(0.311)	(0.022)	(0.022)	(0.022)
Top20Province	–0.293	–0.222	–0.263	–0.219***	–0.229***	–0.236***
	(0.208)	(0.211)	(0.207)	(0.020)	(0.020)	(0.019)
HasAsset	–0.302	–0.341	–0.327	–0.030	–0.162***	–0.051**
	(0.343)	(0.342)	(0.342)	(0.024)	(0.024)	(0.024)
HasLoan	0.015	–0.000	0.020	–0.084***	–0.127***	–0.084***
	(0.193)	(0.195)	(0.194)	(0.027)	(0.027)	(0.027)
NPriorLoan_Applied	0.015***	0.014***	0.014***	–0.096***	–0.098***	–0.070***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Constant	–0.382	–1.499	–0.949	–8.240***	–6.949***	–5.160***
	(1.145)	(1.084)	(1.073)	(0.234)	(0.225)	(0.236)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	1,443	1,443	1,443	397,388	397,388	397,388
Adj. R-squared	0.321	0.320	0.320	0.273	0.305	0.254

Note: Panels A presents the logit regression results with dependent variable Funding Success; this equals 1 if the loan is successfully funded and 0 otherwise. Specifications (1) and (2) focus on the number of total certificates, (3) and (4) focus on the number of important certificates, and (5) and (6) focus on the number of voluntarily applied certificates. Panel B further divides the sample into two groups by credit grade: AA and A, and B and below. Estimated coefficients are reported along with heteroskedasticity robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

7. ROBUSTNESS CHECKS

7.1 Removing Guaranteed Loans

Previous research has shown that deposit insurance reduces depositors' incentive to monitor a bank and encourages risk taking by the secured banks, which poses moral hazard risk (Grossman, 1992; Demirgüç-Kunt and Detragiache, 2002). In the same vein, P2P lenders invest in guaranteed repayment, they lose incentive to prudently screen the quality of the loan applications.

So far, we have shown that the number of certificates is positively related to funding success and loan delinquency, using the full sample, which includes both the secured loans and the unsecured loans. One potential concern of including the secured loans is that the above results may be driven by secured loans whose payments are guaranteed by the platform or backed by collateral, as lenders would be less selective about the loan quality once it is secured. For robustness, we remove the secured loans from the sample, and re-estimate the main results.

Table 7 Panel A presents the number and funding success of loans by type. We can see that the secured loans (i.e. the collateral loans and loans guaranteed by the platform) only make up a small proportion in the full sample, 3.05%, so we expect the secured loan sample to have a low impact on funding success rate. More directly, we re-estimate the main results using a subsample excluding the guaranteed loan, to see if the previous noted relationships still exist. Panel B of Table 7 shows how the number of certificates affects the funding success and the conditional probability of delinquency using the unsecured loans subsample only.

As we can see, the outcomes in Table7 Panel B are qualitatively similar to the full sample results in Table 4 and 6 respectively. Hence, we rule out the concern that the result is merely driven by secured loans and confirm the robustness of findings.

Table 7: Sub Sample Regression: Unsecured Loans

Panel A: Certificates and Funding Success

Guarantee Type	**Full Sample**	**Funded Sample**	**Funding Success Rate** (%)
Guarantee_Credit	605,773	27,136	4.480
Guarantee_Onsite	113,873	113,690	99.839
Guarantee_Collateral	89	87	97.753
Guarantee_Platform	22,557	22,239	98.590
Total	742,292	163,152	21.979

Panel B: Certificates and Delinquency

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Dependent Variable: Hazard Ratio			**Dependent Variable: Funding Success**		
NCertif	1.177***			0.371***		
	(0.005)			(0.004)		
NCertif_Impt		1.237***			0.651***	
		(0.008)			(0.006)	
NCertif_Volun			1.201***			0.467***
			(0.007)			(0.006)

continued on next page

Table 7 *continued*

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Hazard Ratio			Dependent Variable: Funding Success		
CreditGrade	0.420***	0.408***	0.399***	1.648***	1.557***	1.536***
	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.007)
logLoanAmount	1.192***	1.187***	1.193***	–0.880***	–0.912***	–0.873***
	(0.012)	(0.012)	(0.012)	(0.009)	(0.010)	(0.009)
Interest Premium	1.103***	1.090***	1.082***	–0.097***	–0.101***	–0.093***
	(0.009)	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)
Loan Duration (month)	0.972***	0.973***	0.975***	0.021***	0.023***	0.023***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Age	1.004***	1.004***	1.004***	0.036***	0.034***	0.036***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
EduLevel	0.992	0.997	0.993	0.187***	0.208***	0.189***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)
JobIncomeLevel	1.005	1.007**	1.010***	0.325***	0.311***	0.343***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
JobLength	1.039***	1.043***	1.049***	0.135***	0.103***	0.141***
	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.008)
Single_dummy	1.021**	1.023**	1.021**	–0.118***	–0.143***	–0.108***
	(0.010)	(0.010)	(0.010)	(0.018)	(0.018)	(0.018)
Top20Province	1.012	1.012	1.011	–0.192***	–0.192***	–0.199***
	(0.008)	(0.008)	(0.008)	(0.016)	(0.016)	(0.016)
HasAsset	0.985	0.978*	0.981	–0.007	–0.095***	–0.025
	(0.012)	(0.012)	(0.012)	(0.019)	(0.020)	(0.019)
HasLoan	0.923***	0.922***	0.914***	0.045**	–0.002	0.031
	(0.009)	(0.009)	(0.009)	(0.021)	(0.021)	(0.021)
NPriorLoan_Applied	1.005	1.006	1.002	–0.081***	–0.080***	–0.077***
	(0.004)	(0.004)	(0.005)	(0.001)	(0.001)	(0.001)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,309,715	5,309,715	5,309,715	532,982	532,982	532,982
No. of Listings	195,374	195,374	195,374			
No. of Failures	390,964	390,964	390,964			
Pseudo R-squared	0.075	0.075	0.075	0.779	0.788	0.778

Note: Panel A presents the decomposed funding success rate by loans of different types. Guarantee_Credit includes the credit-based loans, Guarantee_Onsite includes loans with onsite authentication by the platform, Guarantee_Collateral includes collateralized loans, and Guarantee_Platform includes platform guaranteed loans.[7] In Panel B, the first three specifications study the how the number of certificates affects the conditional probability (i.e. hazard) of delinquency, where delinquency is defined as default or overdue payments for 1 month or longer. The last three specifications investigate how the number of certificates affects funding success. And the dependent variable Funding Success equals 1 if the loan is successfully funded and 0 otherwise. Specification (1) and (3) focus on the number of total certificates, (2) and (5) focus on the number of important certificates, and (3) and (6) focus on the number of voluntarily applied certificates. Coefficients along with heteroskedasticity robust standard errors in parentheses are reported for logit regressions, Hazard models present hazard ratios along with standard errors clustered at loan level in parentheses clustered at loan level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

[7] According to the information on the platform, credit-based loans (信用认证标 in Chinese) are granted based on the borrower's credit quality, and neither the principal nor the interest are guaranteed. Loans with onsite authentication by the platform (实地认证标 in Chinese) are similar to the credit-based loans with the only difference that the borrowers pass the onsite interview by the platform (or its business partners). Collateralized loans (智能理财标 in Chinese) are granted against the receivables of borrowers, who usually are small business owners. And platform-guaranteed loans (机构担保标 in Chinese) are loans to which the platform (or its business partners) holds joint liability of repayment. Thus the first two types of loans are unsecured and the last two types are secured.

7.2 Principal Components Analysis on Certificate Number

In the models above, we only use on the total number of certificates (or important or voluntarily applied certificates) as our focal variables, without focusing on the combination between certificates of different kinds. The advantage of the simple sum of certificates is that it is intuitive and the coefficients of variables are easy to interpret.

However, assigning each certificate an equal weight may be a questionable assumption. To address this issue, we construct the linear combination of certificates, important certificates, and voluntarily applied certificates using principal components analysis (PCA). And the first principal components (Comp1, Comp1_Impt, and Comp_Volun) are used as the proxies for certificate level.

The eigenvectors and eigenvalues of the principal components for total certificates, important certificates, and voluntarily applied certificates, along with proportions of variation explained by each component are reported in the Internet Appendix 1. And the regression results that related the principal components of certificates to loan performance and funding success are presented in Table 8.

Consistent with our baseline results, the principal component of total, important, and voluntarily applied certificates is positively related to loan delinquency and funding odds, proving the robustness of our main findings.

8. POTENTIAL CHANNELS

So far, we have documented a situation where ineffective certificates distort credit allocation in P2P lending. Specifically, we find that lenders are more willing to invest in listings with higher certificates despite their lower return, poorer credit quality ex ante, and higher delinquency rate ex post.

Biases in cognitive simplification provide a means of understanding the above irrational behavior of lenders. While reasoning by analogy allows lenders to assess the credit quality of borrowers in a simple manner, it is also subject to substantial biases. It is documented that when resorting to analogy, people tend to focus on superficial features without checking if the key underlying assumption is satisfied (Schwenk, 1984, 1988). As a result, the predictability of previous experiences is over-estimated, and representativeness bias is thus introduced into the decision process (Tversky and Kahneman, 1974). More directly, Simon (1959) describes this bias as “the distinction between the objective environment in which the economic actor ‘really’ lives and the subjective environment that he perceives and to which he responds.”

The interaction between lenders and borrowers in P2P lending can be characterized by information asymmetry and adverse selection as documented in Akerlof (1970). Since the seminal works of Jaffee and Russell (1976) and Stiglitz and Weiss (1981), researchers have started to focus on this issue in the credit market. Signaling by borrowers is known to be one of the solutions to alleviate this problem (see Bester, 1985; Besanko and Thakor, 1987; Milde and Riley, 1988). This line of research indicates that high-quality individuals differentiate themselves (i.e. signaling) by choosing an action that cannot be imitated by low-quality individuals.

Table 8: Principal Component Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Hazard Ratio			Dependent Variable: Funding Success		
Comp1	1.055***			0.736***		
	(0.006)			(0.006)		
Comp1_Impt		1.165***			0.811***	
		(0.008)			(0.007)	
Comp1_Volun			1.098***			0.524***
			(0.006)			(0.006)
CreditGrade	0.382***	0.397***	0.394***	1.335***	1.314***	1.554***
	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.007)
logLoanAmount	1.204***	1.214***	1.212***	–0.851***	–0.882***	–0.827***
	(0.013)	(0.014)	(0.013)	(0.009)	(0.009)	(0.009)
Interest Premium	1.076***	1.065***	1.075***	–0.094***	–0.097***	–0.090***
	(0.008)	(0.008)	(0.009)	(0.003)	(0.003)	(0.003)
Loan Duration (month)	0.976***	0.978***	0.976***	0.015***	0.017***	0.015***
	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
Age	1.004***	1.004***	1.004***	0.033***	0.035***	0.034***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
EduLevel	1.001	1.001	1.001	0.215***	0.203***	0.232***
	(0.005)	(0.006)	(0.005)	(0.010)	(0.010)	(0.010)
JobIncomeLevel	1.020***	1.023***	1.022***	0.449***	0.413***	0.418***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
JobLength	1.070***	1.071***	1.070***	0.189***	0.161***	0.201***
	(0.005)	(0.005)	(0.005)	(0.008)	(0.009)	(0.008)
Single_dummy	1.028***	1.030***	1.031***	–0.172***	–0.162***	–0.174***
	(0.010)	(0.010)	(0.010)	(0.018)	(0.018)	(0.018)
Top20Province	1.014*	1.014	1.015*	–0.172***	–0.174***	–0.180***
	(0.008)	(0.009)	(0.008)	(0.016)	(0.016)	(0.016)
HasAsset	0.996	1.002	1.004	0.117***	0.066***	0.116***
	(0.012)	(0.012)	(0.012)	(0.020)	(0.020)	(0.019)
HasLoan	0.910***	0.905***	0.910***	0.131***	0.096***	0.133***
	(0.009)	(0.009)	(0.009)	(0.021)	(0.021)	(0.021)
NPriorLoan_Applied	1.019***	1.021***	1.025***	–0.036***	–0.047***	–0.041***
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,751,091	5,751,091	5,751,091	532,893	532,893	532,893
No. of listings	219,924	219,924	219,924			
No. of failures	398,105	398,105	398,105			
Pseudo R-squared	0.075	0.075	0.075	0.792	0.796	0.781

Note: The first three specifications study the how the principal component of certificates affects the conditional probability (i.e. hazard) of delinquency, where delinquency is defined as default or overdue payments for 1 month or longer. The last three specifications investigate how the principal component of certificates affects funding success. And the dependent variable, Funding Success, equals 1 if the loan is successfully funded and 0 otherwise. Specification (1) and (3) focus on the principal component of total certificates, (2) and (5) focus on the principal component of important certificates, and (3) and (6) focus on the principal component of voluntarily applied certificates. Coefficients along with heteroskedasticity robust standard errors in parentheses are reported for logit regressions, Hazard models present hazard ratios along with standard errors clustered at loan level in parentheses clustered at loan level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

One of the key assumptions shared by the above papers is that the signaling cost is negatively related to an individual's quality (Spence, 1973).[8] In reality, obtaining a certificate on RRD is nearly costless for all applicants, regardless of quality. As a result, certificates are not necessarily associated with high-quality individuals on this P2P platform. Without realizing this subtle difference, bidders reason by analogy and simply interpret the existence of certificates as a positive signal and therefore invest in listings with a higher certificate level.

8.1 Bounded Rationality and Cognitive Simplification

This research documents that inefficient capital allocation outcomes are very different from typical signaling literature and research on certificates. Following the cognitive psychology and behavior economics literature, we show that bounded rationality and biases in cognitive simplification are possible causes for inefficiency.

If certificates were indeed associated with higher credit quality and better loan performance, we could draw a conclusion that certificates serve as positive signals that help high-quality borrowers to attract funds in the credit market, in line with the majority of signaling literature. However, we observe that an increased number of certificates is instead related to poor performance. Therefore it is necessary to explore why lenders are still attracted to listings with a high number of certificates. Previous studies in cognitive psychology and behavioral economics provide a possible explanation.

Making investing decisions on the P2P platform can be complicated, as lenders are exposed to a myriad of information including loan features, applicant's characteristics, investment behavior of other users, certificate level, bidding process, and remaining loan time. Simply understanding these indicators is a time-consuming process, not to mention evaluating borrower's quality, assessing the potential risk and return of each listing, selecting listings that are worth investing in, and allocating appropriate bidding amounts.

To further complicate matters, most lenders on the platform are retail investors with little expertise in finance and investment. Lenders therefore simplify the problem they are faced with by constructing a subjective and simplified world and make decisions accordingly. In decision science literature, this process is also called cognitive simplification (Schwenk, 1984).

Cognitive simplification provides investors with a mental shortcut to assess a borrower's credit quality by relating this new problem to their similar previous experiences, which is known as a representativeness heuristic in cognitive psychology (Tversky, 1974; Tversky and Kahneman, 1974; Kahneman 1991; Kahneman 2003). Compared with formal analysis, this kind of reasoning by analogy is so intuitive and easy that we routinely resort to it in daily decision-making without realizing it (Schunn and Dunbar, 1996; Gentner et al., 2003); even pre-school children are capable of doing so (Holyoak and Thagard, 1997).

There are plenty of real-life situations where certificates are used as a positive signal. For example, higher education level (i.e. an education certificate) reflects better ability (Spence, 1973), and more stars or certificates represent the high standard of a hotel or

---

[8] Another strand of literature also studies costless signaling (i.e. cheap talk) under information asymmetry. For each talk to be effective, additional conditions have to be imposed. For example, Crawford and Sobel (1982) and Austen-Smith and Banks (2000) require similarity in interests among agents, and Franke (1987) proves that the outsider-rationality condition and no-arbitrage condition are needed for costless signaling to function in financial markets. Their assumptions are also violated in the P2P lending market with amateur users, so we do not discuss these papers in detail.

restaurant. Bidders, reasoning by analogy, directly relate certificates to positive attributes without carefully examining if the underlying assumptions are, in fact, true.

Although this phenomenon is rarely mentioned in P2P lending research, a large body of literature in cognitive psychology and decision science has revealed how people reason by analogy when faced with complicated problems, with abundant laboratory and field experience evidence in finance, economics, and management (Gick and Holyoak, 1980; Duhaime and Schwenk, 1985; Schwenk, 1986; Schwenk, 1988; Gavetti et al., 2005; Gavetti and Rivkin, 2005; Gary et al., 2012).

8.2 Difference in Marginal Benefits

Apart from the satisficing decision, another reason that borrowers with lower credit quality are more eager to obtain certificates may be the difference in marginal return. For borrowers, the benefit of having certificates is to boost their credit profiles such that their funding success rate is improved. Given the near zero cost of obtaining certificates, the benefit of certificates is thus an important driving force determining a borrower's certificate level. Therefore, more prominent returns (i.e. lower credit grade) should be associated with higher certificate level.

As shown in Table 6, this benefit from one additional certificate varies dramatically across borrowers in different credit grades. Although one certificate raises funding odds by 44.3% for all the groups, and the increase in HR-rated loans is as large as 109.6%, however, the benefit for the AA-rated group is insignificant.

So, for low-quality borrowers, the marginal benefit is the highest and gives them a high level of incentive to get a high number of certificates. However, high-quality borrowers enjoy little benefit from having more certificates. Hence, this result provides an explanation for the observed adverse selection in certificates, together with borrowers' bounded rationality and the satisficing decision.

8.3 Borrower Myopia and Debt Collection

Certificates may seem costless in terms of time, effort, and monetary expense; however, a large amount of personal information is revealed to the platform via certificates. Although the ramifications of this disclosure may be largely ignored or deemed innocuous during the early stages of RRD use, this disclosure may turn out to be considerably costly later on. In addition to concerns of identity theft, personal information such as address, employer name, and identity of spouse can be used for debt collection purposes by the P2P platform should any default or overdue payments occur.

Following RRD's debt collection policy, in the case of delayed payments, a borrower will first be reminded by SMS and phone call in the first five days. After that, the borrower's designated contact person as recorded on the platform (e.g. relatives, colleagues, employers) will receive notice that the borrower has defaulted on a P2P loan. Should the borrower still refuse to repay, the loan will be transferred to a third-party professional debt collection agency, which will pay the borrower a home visit or even resort to lawsuits.

As a typical borrower normally chooses their certificate level at the beginning stage of borrowing, they may not take into account the full cost of certificates in the debt collection stage if they default. Not realizing the potential costs and being attracted by short-term benefits, borrowers' decision to submit more information to the platform for more certificates may be shortsighted.

We test this hypothesis of myopia by focusing on the different impact of certificates on loan delinquency between borrowers with and without prior default experience on the platform. Empirically, we measure this difference by an interaction term between number of certificates and a dummy variable, Default, which reveal if a borrower has previously defaulted. We argue that borrowers who have personally experienced defaults and the debt collection process should have a better understanding of the indirect cost, which allows us to examine whether the awareness of the long-term cost affects the impact of certificates on delinquent hazard.

The regression results are presented in Table 9. Similar to Table 4, one-month overdue criterion and four-month consecutive overdue criterion are used in Panel A and B. Within each panel, the first, middle, and last two models focus on the number of total, important, and voluntarily applied certificates respectively. The interaction terms are significantly smaller than 1 in both panels across all models, indicating that in impact of certificates in terms of raising delinquency hazard is alleviated if the borrower has default records on the platform. Under the one-month overdue criterion (i.e. in Panel A), although certificates increase delinquency hazards, the influence is weakened by 12.1% (=0.879-1) if the borrower has experienced defaults and the debt collection process.

These findings show that the influence of certificates in terms of affecting delinquency hazard differs significantly between borrowers with different default records. Therefore, when applicants choose their certificate level, the indirect future cost is not taken into consideration, indicating the myopic behavior of borrowers. More importantly, borrower myopia is another important cause for the adverse selection in certificates. The negative relationship between certificates and loan performance is largely reduced (or vanishes) once myopia is corrected.

8.4 Intentional Default

It is possible that certain borrowers do not plan to reply after they receive funding from the platform. For this kind of borrowers, their gain is maximized if they default in the early stage of the loan. Also, they have a stronger incentive to improve their funding success rate by obtaining more certificates, which could result in a positive correlation between the number of certificates and delinquency rate.

Although intentions cannot be observed directly from the data, we can infer intentions based on ex post performance. If a borrower has an early intention of default on a loan that is successfully funded via the platform, to maximize their gain, they should do so immediately after receiving the funds without any repayment, as any payment made to the platform will reduce their return.

We thus define early delinquency as delinquent behavior at the beginning stage of each loan and examine how the number of certificates affects early delinquent behavior. Table 10 shows the variance-corrected multiple-failure Cox proportional hazards model hazard ratios with early delinquent hazard as the dependent variable. While Panel A defines early delinquency as delinquent behavior within the first 3 months after granting of the loan, Panel B recognizes early delinquency as delinquent behavior within the first 1/6 period of the loan. Following the criterion in Table 4, delinquency includes default and any delayed payments. The results are similar when we adopt the four-month consecutive standard, i.e. define delinquency as default or consecutive delayed payments for four months or longer.

Table 9: Alternative Channel: Borrower Myopia

Panel A: One-month Overdue Payment Criterion

Dependent Variable: Default Hazard	(1)	(2)	(3)	(4)	(5)	(6)
NCertif*Default	0.882***	0.879***				
	(0.007)	(0.007)				
NCertif_Impt*Default			0.800***	0.799***		
			(0.008)	(0.008)		
NCertif_Volun*Default					0.785***	0.785***
					(0.008)	(0.008)
NCertif	1.244***	1.240***				
	(0.006)	(0.007)				
NCertif_Impt			1.348***	1.339***		
			(0.011)	(0.011)		
NCertif_Volun					1.319***	1.310***
					(0.010)	(0.010)
Default	6.578***	6.796***	7.650***	7.790***	4.522***	4.561***
	(0.449)	(0.473)	(0.463)	(0.486)	(0.176)	(0.182)
CreditGrade	0.487***	0.503***	0.464***	0.480***	0.450***	0.466***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
logLoanAmount	1.176***	1.167***	1.177***	1.165***	1.187***	1.173***
	(0.014)	(0.013)	(0.014)	(0.014)	(0.015)	(0.014)
Interest Premium	1.069***	1.091***	1.058***	1.081***	1.052***	1.076***
	(0.009)	(0.010)	(0.008)	(0.009)	(0.008)	(0.009)
Loan Duration (month)	0.970***	0.967***	0.970***	0.968***	0.971***	0.969***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age		1.003***		1.003***		1.003***
		(0.000)		(0.000)		(0.000)
EduLevel		1.033***		1.037***		1.034***
		(0.006)		(0.006)		(0.006)
JobIncomeLevel		1.000		1.002		1.003
		(0.004)		(0.004)		(0.004)
JobLength		1.071***		1.070***		1.072***
		(0.005)		(0.005)		(0.005)
Single_dummy		1.013		1.016		1.014
		(0.010)		(0.011)		(0.011)
Top20Province		1.010		1.013		1.014
		(0.009)		(0.009)		(0.009)
HasAsset		0.979*		0.980		0.984
		(0.012)		(0.013)		(0.013)
HasLoan		0.937***		0.939***		0.936***
		(0.009)		(0.009)		(0.009)
NPriorLoan_Applied		1.008**		1.011***		1.010**
		(0.004)		(0.004)		(0.004)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	398,790	398,105	398,790	398,105	398,790	398,105
Pseudo R-square	0.078	0.078	0.078	0.079	0.078	0.078

continued on next page

Table 9 *continued*

Panel B: Four-month Consecutive Overdue Payment Criterion

Dependent Variable: Default Hazard	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NCertif*Default	0.880***	0.871***				
	(0.009)	(0.009)				
NCertif_Impt*Default			0.800***	0.793***		
			(0.010)	(0.011)		
NCertif_Volun*Default					0.780***	0.775***
					(0.011)	(0.011)
NCertif	1.243***	1.247***				
	(0.011)	(0.012)				
NCertif_Impt			1.341***	1.336***		
			(0.016)	(0.017)		
NCertif_Volun					1.319***	1.314***
					(0.015)	(0.015)
Default	30.843***	33.258***	35.655***	36.992***	21.490***	21.795***
	(2.938)	(3.320)	(3.094)	(3.344)	(1.339)	(1.395)
CreditGrade	0.631***	0.657***	0.601***	0.625***	0.585***	0.609***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
logLoanAmount	1.174***	1.155***	1.178***	1.157***	1.184***	1.161***
	(0.015)	(0.014)	(0.016)	(0.015)	(0.017)	(0.016)
Interest Premium	1.050***	1.078***	1.038***	1.068***	1.033***	1.064***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)
Loan Duration (month)	0.980***	0.976***	0.981***	0.977***	0.981***	0.977***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age		1.003***		1.003***		1.003***
		(0.000)		(0.000)		(0.000)
EduLevel		1.037***		1.038***		1.037***
		(0.006)		(0.006)		(0.006)
JobIncomeLevel		1.007*		1.008**		1.010**
		(0.004)		(0.004)		(0.004)
JobLength		1.074***		1.071***		1.073***
		(0.006)		(0.006)		(0.006)
Single_dummy		1.022**		1.024**		1.024**
		(0.011)		(0.011)		(0.011)
Top20Province		1.010		1.013		1.013
		(0.009)		(0.009)		(0.009)
HasAsset		0.986		0.991		0.995
		(0.013)		(0.013)		(0.013)
HasLoan		0.930***		0.932***		0.929***
		(0.010)		(0.010)		(0.010)
NPriorLoan_Applied		1.004		1.011*		1.011**
		(0.006)		(0.006)		(0.005)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	366,223	365,715	366,223	365,715	366,223	365,715
Pseudo R-square	0.073	0.073	0.073	0.073	0.073	0.073

Note: This table investigates the how the number of certificates influences the conditional probability (i.e. hazard) of delinquency differently among borrowers with and without previous default records. In Panel A, delinquency is defined as default or overdue payments for 1 month or longer, while in Panel B, default or overdue payments for 4 months or longer is regarded as delinquency. Within each panel, Specifications (1) and (2) focus on the number of total certificates, (3) and (4) focus on the number of important certificates, and (5) and (6) focus on the number of voluntarily applied certificates. Hazard ratios from the variance-corrected multiple-failure Cox proportional hazards model (Anderson-Gill model) are reported along with standard errors in parentheses clustered at loan level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

Table 10: Intentional Default

Panel A: Delinquent in the First Three Months

Dependent Variable: Default Hazard	(1)	(2)	(3)	(4)	(5)	(6)
NCertif	1.029***	1.051***				
	(0.007)	(0.007)				
NCertif_Impt			1.011	1.029***		
			(0.009)	(0.010)		
NCertif_Volun					0.997	1.020**
					(0.008)	(0.009)
CreditGrade	0.285***	0.289***	0.285***	0.288***	0.285***	0.287***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
logLoanAmount	1.084***	1.052***	1.086***	1.050***	1.089***	1.051***
	(0.016)	(0.019)	(0.016)	(0.019)	(0.016)	(0.019)
Interest Premium	1.051***	1.045***	1.051***	1.046***	1.051***	1.046***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Loan Duration (month)	0.985***	0.985***	0.984***	0.985***	0.984***	0.985***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.002		1.001		1.001
		(0.002)		(0.002)		(0.002)
EduLevel		0.857***		0.862***		0.861***
		(0.012)		(0.012)		(0.012)
JobIncomeLevel		1.019*		1.023**		1.024**
		(0.011)		(0.011)		(0.011)
JobLength		1.033***		1.034***		1.034***
		(0.013)		(0.013)		(0.013)
Single_dummy		0.936***		0.945**		0.942**
		(0.023)		(0.024)		(0.024)
Top20Province		1.070***		1.072***		1.072***
		(0.024)		(0.024)		(0.024)
HasAsset		1.082***		1.090***		1.095***
		(0.029)		(0.029)		(0.029)
HasLoan		0.726***		0.732***		0.733***
		(0.022)		(0.022)		(0.022)
NPriorLoan_Applied		0.983***		0.985***		0.985***
		(0.005)		(0.004)		(0.005)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	8,966	8,886	8,966	8,886	8,966	8,886
Pseudo R-square	0.189	0.192	0.189	0.192	0.189	0.192

continued on next page

Table 10 *continued*

Panel B: Delinquent in the First 1/6 Period

Dependent Variable: Hazard Ratio	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NCertif	1.018*	1.040***				
	(0.011)	(0.011)				
NCertif_Impt			1.021	1.041***		
			(0.014)	(0.015)		
NCertif_Volun					1.002	1.029**
					(0.013)	(0.013)
CreditGrade	0.176***	0.180***	0.176***	0.180***	0.176***	0.180***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
logLoanAmount	1.211***	1.190***	1.210***	1.189***	1.214***	1.190***
	(0.028)	(0.034)	(0.028)	(0.034)	(0.028)	(0.034)
Interest Premium	1.133***	1.124***	1.133***	1.125***	1.133***	1.125***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)
Loan Duration (month)	1.101***	1.100***	1.101***	1.100***	1.101***	1.101***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.005*		1.005*		1.005*
		(0.003)		(0.003)		(0.003)
EduLevel		0.847***		0.850***		0.848***
		(0.016)		(0.016)		(0.016)
JobIncomeLevel		1.008		1.010		1.010
		(0.016)		(0.016)		(0.016)
JobLength		1.019		1.019		1.019
		(0.017)		(0.017)		(0.017)
Single_dummy		0.970		0.974		0.972
		(0.034)		(0.034)		(0.034)
Top20Province		1.102***		1.103***		1.104***
		(0.034)		(0.034)		(0.034)
HasAsset		1.107***		1.106***		1.111***
		(0.041)		(0.041)		(0.042)
HasLoan		0.702***		0.703***		0.704***
		(0.030)		(0.030)		(0.030)
NPriorLoan_Applied		0.975***		0.976***		0.976***
		(0.008)		(0.008)		(0.008)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	5,518	5,496	5,518	5,496	5,518	5,496
Pseudo R-square	0.246	0.248	0.246	0.248	0.246	0.248

Note: This table investigates the impact of certificate number on the conditional probability (i.e. hazard) of early delinquency. In Panel A, early delinquency is defined as default or overdue payments for 1 month or longer within the first three months after the loan is granted, while in Panel B, default or overdue payments for 1 month or longer within the first 1/6 of the loan period is regarded as early delinquency. Within each panel, Specifications (1) and (2) focus on the number of total certificates, (3) and (4) focus on the number of important certificates, and (5) and (6) focus on the number of voluntarily applied certificates. Hazard ratios from the variance-corrected multiple-failure Cox proportional hazards model (Anderson-Gill model) are reported along with standard errors in parentheses clustered at loan level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.

The hazard ratios are significantly larger than 1 for the number of all certificates, important certificates, and voluntary certificates when other factors are controlled, across all versions of definition of early delinquency. These results indicate that intentional default is another reason for the positive relationship between certificates and delinquency. However, the hazard ratios also become much smaller than those in the baseline models of Table 4, which suggests that intentional default is a quantitatively lower order channel.

9. CONCLUSION

Using a large sample of detailed listings and repayment records on one of the PRC's largest P2P platforms, Renrendai, we find that loans from borrowers with higher level of credit certificates have higher delinquency and default rates. However, these loans are also able to attract more funds, garnering a higher funding success rate. This result implies adverse selection in certificates, where borrowers with low credit quality obtain more certificates to boost their credit profile.

This phenomenon is in stark contrast to classical signaling literature, which demonstrates that certificates can serve as a signal that reflects the positive attributes of individuals, thus improving resource allocation efficiency under information asymmetry. Without full awareness of this adverse selection, lenders are still attracted by borrowers with high certificate levels, their despite poorer ex ante quality and higher ex post delinquency. In short, although certificates are widely used as a differentiating signal in resolving adverse selection, they may fail to serve this purpose when the signal itself is inaccurate.

Following previous research in psychology and behavior economics, we propose that differences in marginal benefits from certificates, bounded rationality, cognitive simplification, and borrower myopia account for this phenomenon. Numerous academic studies and real-life experiences demonstrate that high-quality agents differentiate themselves by obtaining certificates. It has become a knee-jerk reaction for both researchers and the general public to interpret certificates as positive signals under information asymmetry.

Our paper provides a counter example to this traditional belief. Although our data is limited to one P2P lending platform in the PRC, the findings may apply to other markets whose participants are amateur and bounded rational. There are many situations where the cost of obtaining certificates is not negatively related to borrower quality, which violates the key assumption of cost signaling. As a result, there is no guarantee that certificates are always associated with positive attributes and favorable outcomes. If signal observers are not sophisticated enough to recognize this nuance, and simply interpret certificates as a positive sign based on cognitive simplification, we will observe similar equilibria in other contexts where low-quality individuals are selected and favored by means of mimicking high-quality individuals.

REFERENCES

Akerlof, George A., 1970. The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics, 84*(3), pp.488–500.

Andersen, P.K. and Gill, R.D., 1982. Cox’s regression model for counting processes: A large sample study. *The Annals of Statistics*, pp.1100–1120.

Aslan, H. and Kumar, P., 2017. Stapled financing, value certification, and lending efficiency. *Journal of Financial and Quantitative Analysis*, *52*(2), pp.677–703.

Auriol, E. and Schilizzi, S. G., 2015. Quality signaling through certification in developing countries. *Journal of Development Economics, 116,* pp.105–121.

Austen-Smith, D. and Banks, J.S., 2000. Cheap talk and burned money. *Journal of Economic Theory*, *91*(1), pp.1–16.

Besanko, D. and Thakor, A.V., 1987. Competitive equilibrium in the credit market under asymmetric information. *Journal of Economic Theory*, *42*(1), pp.167–182.

Best, R. and Zhang, H., 1993. Alternative information sources and the information content of bank loans. *The Journal of Finance*, *48*(4), pp.1507–1522.

Bester, H., 1985. Screening vs. rationing in credit markets with imperfect information. *The American Economic Review*, *75*(4), pp.850–855.

Brach, S., Walsh, G. and Shaw, D., 2018. Sustainable consumption and third-party certification labels: Consumers’ perceptions and reactions. *European Management Journal*, *36*(2), pp.254–265.

Crawford, V.P. and Sobel, J., 1982. Strategic information transmission. *Econometrica, 50*(6), pp.1431–1451.

Demirgüç-Kunt, A. and Detragiache, E., 2002. Does deposit insurance increase banking system stability? An empirical investigation. *Journal of Monetary Economics*, *49*(7), pp.1373–1406.

Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I. and Kammler, J., 2016. Description-text related soft information in peer-to-peer lending–Evidence from two leading European platforms. *Journal of Banking & Finance*, *64*, pp.169–187.

Duarte, J., Siegel, S. and Young, L., 2012. Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies*, *25*(8), pp.2455–2484.

Duhaime, I.M. and Schwenk, C.R., 1985. Conjectures on cognitive simplification in acquisition and divestment decision making. *Academy of Management Review*, *10*(2), pp.287–295.

Focarelli, D., Pozzolo, A.F. and Casolaro, L., 2008. The pricing effect of certification on syndicated loans. *Journal of Monetary Economics*, *55*(2), pp.335–349.

Franke, G., 1987. Costless signalling in financial markets. *The Journal of Finance*, *42*(4), pp.809–822.

Gary, M.S., Wood, R.E. and Pillinger, T., 2012. Enhancing mental models, analogical transfer, and performance in strategic decision making. *Strategic Management Journal*, *33*(11), pp.1229–1246.

Gatti, S., Kleimeier, S., Megginson, W. and Steffanoni, A., 2013. Arranger certification in project finance. *Financial Management*, *42*(1), pp.1–40.

Gavetti, G., Levinthal, D.A. and Rivkin, J.W., 2005. Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), pp.691–712.

Gavetti, G. and Rivkin, J.W., 2005. How strategists really think. *Harvard Business Review*, *83*(4), pp.54–63.

Gentner, D., Loewenstein, J. and Thompson, L., 2003. Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), p.393.

Gick, M.L. and Holyoak, K.J., 1980. Analogical problem solving. *Cognitive Psychology*, *12*(3), pp.306–355.

Gigerenzer, G., 2008. Why heuristics work. *Perspectives on Psychological Science*, *3*(1), pp.20–29.

Godlewski, C.J., 2015. The certification value of private debt renegotiation and the design of financial contracts: Empirical evidence from Europe. *Journal of Banking & Finance*, *53*, pp.1–17.

Godlewski, C.J. and Sanditov, B., 2018. Financial institutions network and the certification value of bank loans. *Financial Management*, *47*(2), pp.253–283.

Grossman, R.S., 1992. Deposit insurance, regulation, and moral hazard in the thrift industry: Evidence from the 1930's. *The American Economic Review*, 82(4), pp.800–821.

He, Q., Li, X. and Zhu, W., 2019. Political connection and the walking dead: Evidence from China's privately owned firms. *International Review of Economics & Finance.*

Holyoak, K.J. and Thagard, P., 1997. The analogical mind. *American Psychologist*, *52*(1), p.35.

Iyer, R., Khwaja, A., Luttmer, E. and Shue, K., 2009. Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending?. *SSRN Electronic Journal*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1570115.

Jaffee, D. M. and Russell, T., 1976. Imperfect information, uncertainty, and credit rationing. *The Quarterly Journal of Economics*, *90*(4), pp.651–666.

James, C., 1987. Some evidence on the uniqueness of bank loans. *Journal of Financial Economics*, *19*(2), pp.217–235.

Kahneman, D., 1991. Article commentary: Judgment and decision making: A personal view. *Psychological Science*, *2*(3), pp.142–145.

Kahneman, D., 2003. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*(5), pp.1449–1475.

Lin, M., Prabhala, N.R. and 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), pp.17–35.

Liu, D., Brass, D. J., Lu, Y., and Chen, D., 2015. Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding. *Mis Quarterly, 39*(3), pp.729–742.

Lummer, S.L. and McConnell, J.J., 1989. Further evidence on the bank lending process and the capital-market response to bank loan agreements. *Journal of Financial Economics*, *25*(1), pp.99–122.

Michels, J., 2012. Do unverifiable disclosures matter? Evidence from peer-to-peer lending. *The Accounting Review*, *87*(4), pp.1385–1413.

Milde, H. and Riley, J.G., 1988. Signaling in credit markets. *The Quarterly Journal of Economics*, *103*(1), pp.101–129.

Milgrom, P. R. 1981. Good news and bad news: Representation theorems and applications. The Bell Journal of Economics 12 (2): 380–391.

Riley, J. G., 1979. Informational equilibrium. *Econometrica, 47*(2), pp.331–359.

Schunn, C.D. and Dunbar, K., 1996. Priming, analogy, and awareness in complex reasoning. *Memory & Cognition*, *24*(3), pp.271–284.

Schwenk, C.R., 1984. Cognitive simplification processes in strategic decision making. *Strategic Management Journal, 5*(2), pp.111–128.

Schwenk, C.R., 1986. Information, cognitive biases, and commitment to a course of action. *Academy of Management Review*, *11*(2), pp.298–310.

Schwenk, C.R., 1988. The cognitive perspective on strategic decision making. *Journal of Management Studies*, *25*(1), pp.41–55.

Simon, H.A., 1955. A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), pp.99–118.

Simon, H.A., 1959. Theories of decision-making in economics and behavioral science. *The American Economic Review*, *49*(3), pp.253–283.

Simon, H.A., 1979. Rational decision making in business organizations. *The American Economic Review*, *69*(4), pp.493–513.

Spence, M., 1973. Job market signaling. *Quarterly Journal of Economics, 87*(3), pp.355–374.

Stiglitz, J.E. and Weiss, A., 1981. Credit rationing in markets with imperfect information. *The American Economic Review*, *71*(3), pp.393–410.

Tversky, A., 1974. Assessing uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.148–159.

Tversky, A. and Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), pp.1124–1131.

Zhang, J. and Liu, P., 2012. Rational herding in microloan markets. *Management Science*, *58*(5), pp.892–912.

APPENDIX 1: VARIABLE DEFINITION

Variable	Definition
Borrower Characteristics	
NCertif	The number of total certificates.
NCertif_Impt	The number of important certificates.
NCertif_Volun	The number of voluntarily applied certificates.
NCertif*Default	The interaction between NCertif and Defaulted.
NCertif_Impt*Default	The interaction between NCertif_Impt and Defaulted.
NCertif_Volun*Default	The interaction between NCertif_Volun and Defaulted.
CreditGrade	Credit grade assigned by the platform, including seven grades AA, A, B, C, D, E, and HR. AA equals 7; A equals 6; B equals 5; C equals 4; D equals 3; E equals 2; and HR equals 1.
Age	Age of each borrower.
EduLevel	Education level. Equals 4 if the borrower's highest qualification is a master's degree or above; 3 if the borrower's highest qualification is a bachelor's degree; 2 if the borrower's highest qualification is post-tertiary; and 1 if the borrower's highest qualification is secondary or below.
JobIncomeLevel	Monthly income level. 7 means more than 50,000 RMB; 6 means between 20,000 and 50,000 RMB; 5 means between 10,000 and 20,000 RMB; 4 means between 5,000 and 10,000 RMB; 3 means between 2,000 and 5,000 RMB; 2 means between 1,000 and 2,000 RMB; and 1 means less than 1,000 RMB.
JobLength	Employment length in years. 4 indicates more than 5 years; 3 means between 3 and 5 years; 2 indicates between 1 and 3 years; and 1 indicates less than 1 year.
Single_dummy	Dummy variable that equals 1 if the marital status is single; and 0 otherwise.
Top20Province	Dummy variable that equals 1 if the borrower is from the top 20 provinces; and 0 otherwise.
HasAsset	Dummy variable that equals 1 if the borrower owns house or car; and 0 otherwise.
HasLoan	Dummy variable that equals 1 if the borrower has car loan or mortgage loan; and 0 otherwise.
NPriorLoan_Applied	Number of prior applied loans of each borrower.
Default	Dummy variable that equals 1 if the borrower has defaulted a loan on RRD.com and 0 otherwise.
Comp1	First principal component of total certificates.
Comp1_Impt	First principal component of important certificates.
Comp1_Volun	First principal component of voluntary certificates.
Loan Contract Terms	
Loan Amount (k)	Requested loan amount in thousands of RMB of each loan.
Interest Premium	Difference between the loan rate and the corresponding benchmark interest rate of each loan.
Loan Duration (month)	Loan duration in months of each loan.
Interest Rate	Loan interest rate of each loan.
Loan Performance	
BadDebt	An ordered discrete variable which equals 0 if the loan is repaid on time in each period, 1 if the loan is completely repaid but with overdue records, and 2 if the loan is unrepaid.
BadDebt (=0)	Dummy variable that equals 1 if the loan is repaid on time in each period; and 0 otherwise.

BadDebt (=1)	Dummy variable that equals 1 if the loan is completely repaid but with overdue records; and 0 otherwise.
BadDebt (=2)	Dummy variable that equals 1 if the loan is unrepaid; and 0 otherwise.

INTERNET APPENDIX 1: EIGENVALUES AND EIGENVECTORS OF PRINCIPAL COMPONENTS

Panel A: Proportion of First Component (Comp1)

	Eigenvalue	**Proportion** (%)
Comp1	3.668	17.47
Comp1_Impt	3.555	29.63
Comp1_Volun	2.915	18.22

Panel B: Eigenvector of First Component (Comp1)

Certificate	**Comp1**	**Comp1_Impt**	**Comp1_Volun**
Onsite Authentication	0.381	0.377	0.343
Property Ownership	–0.024	0.018	–0.035
Loan Purpose	–0.012		–0.015
Remote Video	–0.021	0.013	–0.040
Bank Statement (Salary)	0.456	0.469	0.563
Credit Report	0.401	0.418	0.490
Other	–0.001		–0.002
ID Number	0.020	0.016	
Platform Training	–0.139		
Social Network	–0.018		–0.028
Phone Bill	–0.027	0.011	–0.038
Child	–0.011		–0.017
ID Card	–0.167	0.320	
Microblog	0.280		–0.032
Residence Proof	–0.024	0.013	–0.036
Occupation	–0.026	0.468	0.561
Academic Qualification	0.454		–0.018
Mobile Phone	–0.017	–0.374	
Marriage Certificate	–0.386		–0.042
Car	–0.027	0.019	–0.034

Note: Panel A presents the eigenvalues of the principal components for total certificates, important certificates, and voluntarily applied certificates, along with proportions of variation explained by each component. Panel B reports the corresponding eigenvector of each component. The definitions are presented in Appendix 1.

INTERNET APPENDIX 2: ROBUSTNESS TESTS USING DIFFERENT LENGTHS OF OVERDUE PAYMENTS

Panel A: One-month Overdue Payment Criterion (Single-Failure)

Dependent Variable: Delinquent Hazard	(1)	(2)	(3)	(4)	(5)	(6)
NCertif	1.097***	1.090***				
	(0.005)	(0.005)				
NCertif_Impt			1.117***	1.106***		
			(0.007)	(0.007)		
NCertif_Volun					1.112***	1.102***
					(0.007)	(0.007)
CreditGrade	0.394***	0.404***	0.386***	0.397***	0.381***	0.392***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
logLoanAmount	1.216***	1.199***	1.215***	1.196***	1.220***	1.200***
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.015)
Interest Premium	1.080***	1.099***	1.072***	1.091***	1.072***	1.092***
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.009)
Loan Duration (month)	0.973***	0.973***	0.974***	0.974***	0.974***	0.974***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.007***		1.006***		1.006***
		(0.001)		(0.001)		(0.001)
EduLevel		1.015*		1.018**		1.015*
		(0.008)		(0.008)		(0.008)
JobIncomeLevel		1.004		1.006		1.007
		(0.006)		(0.006)		(0.006)
JobLength		1.080***		1.079***		1.082***
		(0.008)		(0.008)		(0.008)
Single_dummy		1.013		1.015		1.013
		(0.012)		(0.012)		(0.012)
Top20Province		1.024**		1.025**		1.025**
		(0.011)		(0.011)		(0.011)
HasAsset		1.012		1.012		1.012
		(0.018)		(0.018)		(0.018)
HasLoan		0.923***		0.923***		0.920***
		(0.012)		(0.012)		(0.012)
NPriorLoan_Applied		1.014***		1.015***		1.013***
		(0.003)		(0.004)		(0.004)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,571,323	5,565,412	5,571,323	5,565,412	5,571,323	5,565,412
No. of Listings	220,612	219,930	220,612	219,930	220,612	219,930
No. of Failures	53,923	53,784	53,923	53,784	53,923	53,784
Pseudo R-square	0.079	0.079	0.079	0.079	0.079	0.079

continued on next page

Internet Appendix 2 *table continued*

Panel B: Four-month Consecutive Overdue Payment Criterion (Single-Failure)

Dependent Variable: Delinquent Hazard	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NCertif	1.057***	1.056***				
	(0.005)	(0.005)				
NCertif_Impt			1.079***	1.075***		
			(0.007)	(0.007)		
NCertif_Volun					1.068***	1.066***
					(0.007)	(0.007)
CreditGrade	0.434***	0.447***	0.429***	0.441***	0.425***	0.437***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
logLoanAmount	1.275***	1.247***	1.275***	1.246***	1.278***	1.248***
	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)
Interest Premium	1.087***	1.118***	1.082***	1.113***	1.080***	1.110***
	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)
Loan Duration (month)	0.992***	0.990***	0.992***	0.991***	0.993***	0.991***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.008***		1.008***		1.008***
		(0.001)		(0.001)		(0.001)
EduLevel		0.998		0.999		0.998
		(0.007)		(0.007)		(0.007)
JobIncomeLevel		1.017***		1.018***		1.019***
		(0.006)		(0.006)		(0.006)
JobLength		1.068***		1.066***		1.067***
		(0.008)		(0.008)		(0.008)
Single_dummy		1.042***		1.042***		1.042***
		(0.012)		(0.012)		(0.012)
Top20Province		1.032***		1.032***		1.032***
		(0.011)		(0.011)		(0.011)
HasAsset		0.991		0.991		0.992
		(0.017)		(0.017)		(0.017)
HasLoan		0.915***		0.915***		0.913***
		(0.013)		(0.013)		(0.013)
NPriorLoan_Applied		1.005		1.005		1.003
		(0.006)		(0.006)		(0.006)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,643,990	5,637,598	5,643,990	5,637,598	5,643,990	5,637,598
No. of Listings	220,612	219,930	220,612	219,930	220,612	219,930
No. of Failures	46,560	46,496	46,560	46,496	46,560	46,496
Pseudo R-square	0.063	0.064	0.063	0.064	0.063	0.064

continued on next page

Internet Appendix 2 *table continued*

Panel C: Two-month Consecutive Overdue Payment Criterion (Multiple-Failure)

Dependent Variable: Delinquent Hazard	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NCertif	1.154***	1.151***				
	(0.006)	(0.006)				
NCertif_Impt			1.218***	1.212***		
			(0.008)	(0.008)		
NCertif_Volun					1.182***	1.180***
					(0.008)	(0.008)
CreditGrade	0.399***	0.410***	0.389***	0.399***	0.380***	0.391***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
logLoanAmount	1.204***	1.197***	1.202***	1.194***	1.210***	1.199***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)
Interest Premium	1.082***	1.106***	1.069***	1.094***	1.060***	1.085***
	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)
Loan Duration (month)	0.984***	0.981***	0.982***	0.982***	0.986***	0.983***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age		1.004***		1.004***		1.004***
		(0.000)		(0.000)		(0.000)
EduLevel		0.990*		0.993		0.990*
		(0.005)		(0.005)		(0.005)
JobIncomeLevel		1.012***		1.014***		1.016***
		(0.004)		(0.004)		(0.004)
JobLength		1.058***		1.056***		1.059***
		(0.005)		(0.005)		(0.005)
Single_dummy		1.026***		1.027***		1.026***
		(0.010)		(0.010)		(0.010)
Top20Province		1.012		1.013		1.013
		(0.008)		(0.008)		(0.008)
HasAsset		0.980*		0.977**		0.980
		(0.012)		(0.012)		(0.012)
HasLoan		0.908***		0.908***		0.903***
		(0.009)		(0.009)		(0.009)
NPriorLoan_Applied		1.003		1.004		0.999
		(0.005)		(0.006)		(0.006)
Yr Qtr FE	YES	YES	YES	YES	YES	YES
Observations	5,757,306	5,751,091	5,757,306	5,751,091	5,757,306	5,751,091
No. of Listings	220,606	219,924	220,606	219,924	220,606	219,924
No. of Failures	378,150	377,586	378,150	377,586	378,150	377,586
Pseudo R-square	0.071	0.071	0.071	0.071	0.070	0.071

Note: We examine the robustness of our findings by using alternative estimation technique and definition of delinquency. Panels A and B repeat the estimation of Table 4 using the single-failure model, where each loan observation after the first delinquency is ignored. In Panel C, we change the definition of delinquency to default or consecutive overdue for 2 months or longer, and re-estimate the multiple-failure model. Hazard ratios from the variance-corrected multiple-failure Cox proportional hazards model (Anderson-Gill model) are reported along with standard errors in parentheses clustered at listing level. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. The definitions of all variables are presented in Appendix 1.